Dienst Landbouwkundig Onderzoek

Instituut voor Agrotechnologisch Onderzoek

Voorspelling van de productkwaliteit in relatie tot grondstofkarakteristieken

A. Braaksma
E.C. Wilkinson
J. Oosterhaven
P.C.M. van Eijck
E.S.A. Biekman
F.A.J. Bossers

Vertrouwelijk

ato-dlo

Rapport B301 / Januari 1998

ato-dlo

Voorspelling van productkwaliteit in relatie tot grondstofkarakteristieken

-statistische relaties over de periode 1993-1997

Kleur
Grauw
Textuur

VERTROUWELIJK

Januari 1998

A.Braaksma
E.C.Wilkinson
J. Oosterhaven
P.C.M. van Eijck
E.S.A. Biekman
F.A.J. Bossers

Agrotechnologisch Onderzoek Instituut (ATO-DLO)
Bornsesteeg 59
Postbus 17
6700 AA
Wageningen
tel. 0317-475000
fax. 0317-475347

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

225119

Voorwoord

Het onderzoek waarvan hier verslag wordt gedaan, richtte zich op het vinden van statistische relaties die gebruikt kunnen worden om de producteigenschappen ***bakkleur, grauwverkleuring en textuuraspecten*** op voorhand te voorspellen door middel van bepalingen aan de rauwe aardappel op partijniveau.
Daartoe zijn in de periode 1993 - 1997 in totaal 100 partijen aardappelen uit de praktijk verwerkt op de ATO pilotlijn, waaronder een aantal controle-partijen. Hierbij zijn tientallen grondstofkarakteristieken bepaald. Al deze gegevens zijn statistisch op hun samenhang met het gewenste productkenmerk geanalyseerd en de uitkomsten ervan staan in dit verslag vermeld.

Het (achtergrond)onderzoek rondom het grondstofproject wordt in een afzonderlijke rapportage vermeld.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Inhoudsopgave

1. INLEIDING **4**

2. MATERIAAL EN METHODEN **5**

2.1 Uitgangsmateriaal **5**

2.2 Statistische methoden **5**

2.3 Heterogeniteit en variatie-bronnen 6

2.4 Belangrijkste resultaten **8**

3. RESULTATEN **10**

3.1 Bakkleur **10**
3.1.1 Inleiding 10
3.1.2 Resultaten 10
3.1.3 Conclusies 14

3.2 Grauwverkleuring **14**
3.2.1 Grauwverkleuring van aardappelen 14
3.2.2 Grauwverkleuring en selectie van grondstofeigenschappen 15
3.2.3 Statistische selectie van relevante componenten en modelvorming 17
3.2.4 Algemeen Grauw-model 18
3.2.5 Effect van bewaring 19
3.2.6 Effect van oogstseizoen 19
3.2.7 Ras: Agria versus Bintje 21
3.2.8 Conclusie 22

3.3 Textuur **22**
3.3.1 Inleiding 22
3.3.2 Methoden 23
3.3.3 Resultaten 23
3.3.4 Voorspellende modellen 27
3.3.5 Onderwatergewicht versus drogestofgehalte 30
3.3.6 Conclusie 31

4. DISCUSSIE **32**

BIJLAGE 1 TOELICHTING STATISTISCHE ANALYSE **35**

BIJLAGE 2 VARIANTIE-ANALYSE VAN DE GEMETEN GRONDSTOF- EN PRODUCTEIGEN-SCHAPPEN **39**

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

1. Inleiding

Het is bekend van een aantal kwaliteitsaspecten van aardappelproducten dat deze zeer sterk worden bepaald door chemische en/of fysische kenmerken van de rauwe aardappel.
Voor de eindkwaliteit van gefrituurde aardappelproducten, in termen van kleur, grauwverkleuring en textuureigenschappen, is deze samenhang ook bekend of wordt op zijn minst vermoed aanwezig te zijn. Doelstelling van het project is om aan de hand van de grondstofsamenstelling van het rauwe produkt, een voorspelling van de kwaliteitseigenschappen na verwerking te geven. Dit moet resulteren in meetmethoden die in staat zijn de kleur, de mate van grauwverkleuring en belangrijke textuur-eigenschappen te voorspellen.

De proefopzet was zodanig dat er jaarlijks 20-30 aardappelpartijen op de pilotlijn zijn verwerkt in de periode 1993 - 1997. Belangrijk uitgangspunt is geweest om een breed scala van rassen en partijen te kiezen waardoor er veel variatie in samenstelling en in productkwaliteit is geïntroduceerd. Van alle partijen zijn vele chemische en fysische karakteristieken bepaald. De verwerkingscondities zijn gedurende het gehele project constant gehouden en na de verwerking zijn de relevante kwaliteitseigenschappen vastgesteld middels sensorische en/of objectieve meetmethoden. Dit heeft een grote database met gegevens opgeleverd die het mogelijk maakt om de grondstof-samenstelling te relateren aan de produktkwaliteit. Dit verslag beschrijft enkel de uitkomsten van de statistische analyses. In een ander verslag komt het achtergrondsonderzoek aan bod.

Tabel 1. Overzicht van de partijen die statistisch geanalyseerd zijn.

ras	aantal partijen	ras	aantal partijen
Agria	15	Disco	3
Arcade	1	Felsina	2
Asterix	9	Irene	2
Aziza	4	Nicola	5
Bintje	17	Saturna	2
Bildtstar	1	Santana	2
		Turbo	10

Doel van de statistische analyse is te komen tot inzicht in die factoren die de verwerkingskwaliteit in belangrijke mate beïnvloeden om daarmee te komen tot goede voorspellingsmogelijkheden.
Het eerste deel van het verslag geeft een overzicht van de gevolgde statistische procedures, gevolgd door de resultaten en de interpretatie wat betreft kleur, grauwverkleuring en textuur. Het verslag eindigt met een discussie en met de richting die het (achtergrond)onderzoek moet nemen om te komen tot eenvoudige, snelle en robuuste meetmethoden die betrouwbaar de kwaliteit voorspellen.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

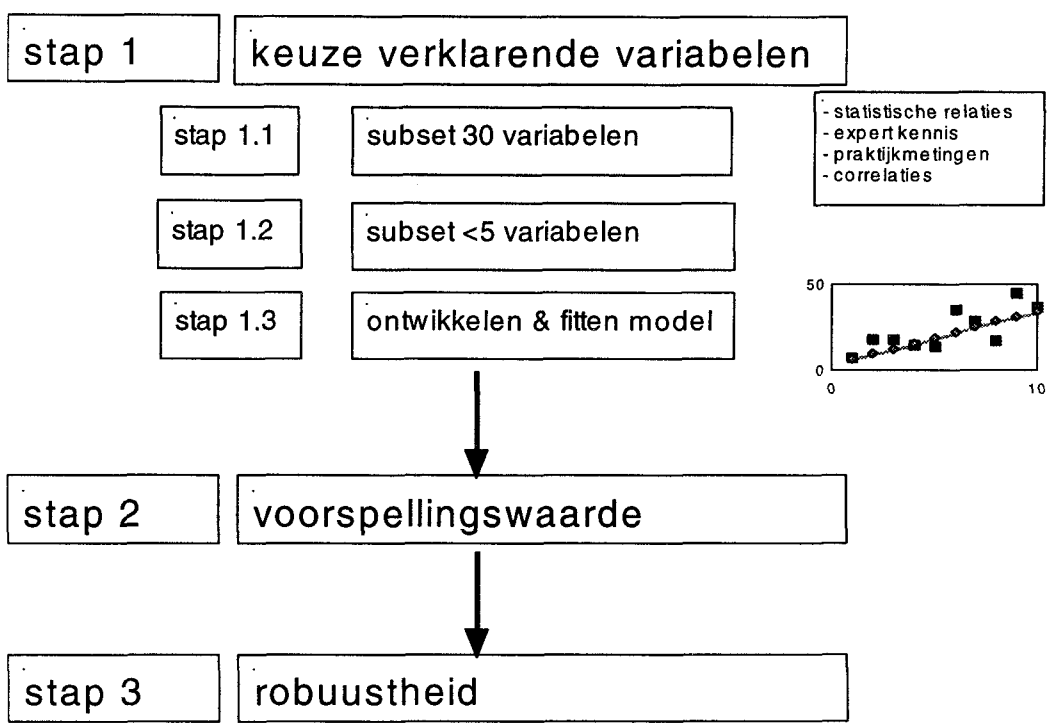
2. Materiaal en Methoden

2.1 Uitgangsmateriaal

In de periode 1993-1997 zijn er in 7 verwerkingsrondes in totaal 100 partijen aardappelen verwerkt, waarvan 79 in de statistische analyse zijn betrokken. De aardappelen zijn bemonsterd en geanalyseerd volgens de procedures zoals die beschreven staan in het rapport: *Methoden voor de karakterisering van de grondstof Aardappel* (versie van januari 1998). Het resultaat van de analyses aan de ruwe grondstof en aan het eindproduct is een complete dataset welke aan diverse statistische analyses is onderworpen.

2.2 Statistische methoden

Na elke verwerking is een statistische analyse uitgevoerd van de verzamelde gegevens tot en met de betreffende ronde. Deze analyse bestond steeds uit twee delen. Enerzijds werd een overzicht gemaakt van de correlaties tussen de grondstofeigenschappen onderling en tussen de producteigenschappen onderling. Daarnaast werden steeds voorspellingsmodellen ontwikkeld die de grondstofeigenschappen relateren aan de belangrijkste producteigenschappen. Er zijn ongeveer 60 grondstofeigenschappen en ongeveer 40 producteigenschappen gemeten. Deze grote hoeveelheid variabelen eist een speciale statistische aanpak. Om het verslag leesbaar te houden staat in bijlage 1 een gedetailleerd overzicht van de gebruikte statistiek. De aanpak voor de modellering staat hieronder schematisch weergegeven:



Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

2.3 Heterogeniteit en variatie-bronnen

Inleiding
Een belangrijke beperkende factor zowel bij het ontwikkelen als bij het gebruiken van de modellen, is de *nauwkeurigheid* waarmee men de grondstofeigenschap cq. producteigenschap kan bepalen. Bij het doen van correcte voorspellingen op *partij*-niveau dan zijn er meerdere bronnen van variatie relevant die het resultaat kunnen beïnvloeden, nl.;

- verschillen tussen knollen/frites;
- verschillen binnen knollen/frites;
- verwerkingseffecten;
- meetfout;
- bij sensorische waarnemingen, verschillen tussen panelleden/productexperts.

In de vijfde verwerkingsronde is speciaal aandacht gegeven aan de bronnen van variatie. Er is voor een proefopzet gekozen waardoor de relatieve bijdrage van de verschillende bronnen van variatie in kaart kon worden gebracht. Hiermee zijn we in staat om advies te geven over de optimale monstergrootte en meetmethode.

Proefopzet en analyse van de resultaten van de vijfde verwerkingsronde
De proefopzet voor de vijfde ronde is speciaal ontworpen om informatie over variatie te achterhalen. Er werden in totaal negen verschillende partijen gebruikt. Vijf van de negen partijen werden bij aankomst in twee batches opgesplitst, die verder werden behandeld alsof ze twee aparte partijen waren. De onderzoekers die de analyses uitvoerden wisten niet welke batches bij elkaar hoorden.

Tabel 2: Proefopzet vijfde ronde

Partij	batch-code	maat	Partij	batch-code	maat
Agria 1	AG01	alles		BI03	klein
	AG03	alles		BI02	midden
	AG04	klein		BI06	groot
	AG06	midden	Bintje 2	BI04	alles
	AG05	groot	Bildstar	BS01	alles
Agria 2	AG02	alles	Nicola	NI01	alles
Asterix	AS01	alles	Saturna	SA01	alles
	AS02	alles		SA02	alles
Bintje 1	BI01	alles	Turbo	TU01	alles
	BI05	alles		TU02	alles

Het verschil in gemeten waarde tussen twee batches van dezelfde partij vormt dus een objectieve maat voor de nauwkeurigheid van de meetmethode. Ook werden één partij Agria en één partij Bintje in drie maatsorteringen (klein, midden, groot) opgesplitst, die verder werden behandeld alsof ze aparte partijen waren. Kort samengevat, hoewel bij de vijfde ronde maar negen partijen binnen kwamen, werden ze zodanig opgesplitst dat

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

uiteindelijk 20 "partijen" werden geanalyseerd. Tabel 2 laat zien hoe de partijen werden opgesplitst.

Analysemethodes
Van elke analysemethode is de meetfout bekend, waarmee de nauwkeurigheid van elke bepaling is vastgelegd. Zo is voor de chemische analyses aan de grondstof de fout <5%. De uiteindelijke nauwkeurigheid van de bepaling van het partij-gemiddelde is ook afhankelijk van de monstergrootte. Een grotere steekproef leidt immers tot een nauwkeuriger bepaling, maar het maakt ook uit of je aan één monster meerdere keren meet of aan meerdere monsters slechts één keer meet. Dit laatste kan met name bij sensorische bepalingen het geval zijn.
Is een bepaald kenmerk sterk variabel in een partij, dan kan dit kenmerk wel nauwkeurig worden gemeten met de standaardmonstername van 20 knollen, maar het karakteriseert de partij niet; een tweede monster geeft dan sterk afwijkende gehaltes. De ene partij is op grond hiervan dan niet van een andere partij te onderscheiden.
Voor de analyse van deze variaties en het onderscheidingsvermogen van de bepalingen wordt een *variantiecomponentenanalyse* gebruikt. Deze methode levert een schatting op van de zogenaamde *variantiecomponenten*, die een maat zijn voor de relatieve bijdrage van de verschillende variatiebronnen. Waar mogelijk werden de oorspronkelijke gegevens gebruikt, dat wil zeggen de duplo metingen, de beoordelingen van individuele panelleden/productexperts enz. Op deze manier kon de relatieve bijdrage van de meetfout, monstergrootte, verwerking en andere factoren aan de totale variatie worden achterhaald.
Met de geschatte variantiecomponenten kunnen we dan berekenen wat de nauwkeurigheid is van een meetmethode zoals die in het VAVI-onderzoek is uitgevoerd. Maar we kunnen ook "what if" analyses doen, waarin we het effect zien van meer/minder metingen, meer/minder monstertrekkingen meer/minder panelleden enz.
De nauwkeurigheid van de schatting van het partij-gemiddelde bij een gegeven meet-scenario wordt gegeven door de variantie van de schatting. Hoe kleiner de variantie, hoe nauwkeuriger de schatting. Om te kunnen zien of dit getal voldoende klein is, moet men de variantie aan andere maten voor dezelfde eigenschap relateren. Er zijn hier twee indicatoren gebruikt:

1. *Percentage fout*
Deze maat geeft de standaard fout (wortel van de variantie) van de schatting als percentage van het gemiddelde (het algemene gemiddelde voor de vijfde ronde wordt gebruikt). Hoe kleiner dit getal, hoe nauwkeuriger de schatting. Deze maat is niet relevant voor de sensorische bepalingen, waar de schaal en dus de gemiddelde waarde volstrekt willekeurig is.

2. *Discriminerend vermogen*
Deze maat geeft de standaard fout van de schatting als percentage van de totale range van waarden (berekend over alle vijf rondes). Immers, het percentage fout kan wel klein zijn, maar als de verschillen tussen de partijen ook klein zijn, dan is het voor de doeleinden van dit projekt toch niet zo'n aantrekkelijke eigenschap. Een lage waarde voor deze maat betekent dat de meetmethode in staat is goed tussen de partijen te kunnen discrimineren.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

2.4 Belangrijkste resultaten

Nauwkeurigheid metingen bij huidige meet-scenario
Een overzicht van zowel het percentage fout als het discriminerend vermogen is te zien in bijlage 2. De analyseresultaten laten duidelijk zien welke eigenschappen goed of minder goed te meten zijn. Idealiter wil men zowel een lage percentage fout als een goed discriminerend vermogen. Het onderwatergewicht (OWG) van ongeschilde aardappels is hier een voorbeeld van. Aan de andere kant heeft kalium (code 110[2]) weliswaar een lage meetfout, maar de variatie binnen een partij is zo groot dat het een slecht discriminerend vermogen heeft (zie bijlage 2). Toch kunnen we concluderen dat de meeste eigenschappen nauwkeurig te meten zijn.
Dat een lage percentage fout niet noodzakelijk een goed discriminerend vermogen hoeft te betekenen, is duidelijk te zien aan de resultaten voor citroenzuur (112[2]) en appelzuur (112[3]). Citroenzuur heeft een wat lagere percentage fout maar een veel hogere waarde voor het discriminerend vermogen, wat het een minder aantrekkelijke eigenschap maakt dan appelzuur voor gebruik in voorspellingsmodellen.

"What if" analyses
De variantiecomponenten-analyse stelt ons in staat om "what if" analyses uit te voeren. Hierdoor krijgen we een indruk van de verandering in meetnauwkeurigheid als we andere meet-scenario's kiezen. Enkele voorbeelden voor een drietal eigenschappen zijn te zien in tabel 3 waarbij het huidige scenario wordt weergegeven met een donkere achtergrond.
We hebben al gezien dat OWG nauwkeurig te meten is. De "what if" analyse laat zien dat zelfs het weglaten van de duplo-meting geen grote gevolgen zou hebben voor het discriminerend vermogen. Blijkbaar is de monstergrootte voldoende groot om de partij afdoende te karakteriseren.
Voor de sensorische eigenschappen kan men zowel het aantal herhalingen variëren als het aantal experts/panelleden. Dat deze keuze gevolgen kan hebben voor de meetnauwkeurigheid is aan de resultaten voor TXA te zien. De huidige scenario van twee keer afbakken en twee experts levert 4 bepalingen op en een discriminerend vermogen van 8.8%. Eén keer afbakken en vier experts zou ook 4 bepalingen geven, maar dan met een discriminerend vermogen van 11.6%. Het is duidelijk dat men beter kan kiezen om vaker af te bakken dan om meer experts erbij te betrekken.
Voor de eigenschap krokante korst hebben de optie "weglaten van een herhaling" en de optie "halveren van het aantal panelleden" nagenoeg hetzelfde effect (een discriminerend vermogen van 9.3% resp. 8.98%). De uiteindelijke keuze zal dan afhankelijk zijn van de verwachte besparing in inspanning en kosten.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Tabel 3: "What if" analyse: effect verschillende meetmethode scenario's op discriminerend vermogen. Huidige scenario wordt weergegeven met donkere achtergrond.

	Aantal metingen		discriminerend vermogen
	herhaling	experts/panelleden	s.e./range (%)
OWG ongeschild	2	n.v.t.	1.6
	1	n.v.t.	2.1
TXA	2	2	8.8
nat/droog	3	2	7.2
(experts)	2	4	8.2
	3	1	8.2
	1	4	11.6
krokante korst (sensorisch analytisch panel)	3	20	7.6
	3	10	8.9
	2	20	9.3
	2	10	10.9

Conclusies

We hebben de relatieve bijdrage van verschillende variatiebronnen in kaart gebracht voor een groot aantal eigenschappen zowel van de aardappel als van de friet. Hiermee kunnen we de nauwkeurigheid van de huidige meetmethode bepalen, en advies geven over een optimaal meetmethode-scenario. Deze resultaten zijn relevant niet alleen binnen het kader van dit projekt maar ook daarbuiten.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

3. Resultaten

3.1 Bakkleur

3.1.1 Inleiding

Voor de praktijk is het van groot belang om op voorhand van een partij aangeleverde aardappelen de uiteindelijke bakkleur van de frites te kunnen inschatten, daar de bakkleur van de frites nog altijd als één van de belangrijkste kwaliteitskenmerken wordt beschouwd.
De kleur van frites wordt voor een belangrijk deel bepaald door de concentratie reducerende suikers in het weefsel. Er zijn echter partijen en seizoenen waarbij de relatie niet goed opgaat. Doel van deze analyse is dan ook om een betere voorspelbaarheid te verkrijgen van de kleur van het eindproduct op basis van grondstofeigenschappen.

3.1.2 Resultaten

Allereerst is een selectie gemaakt uit alle parameters om binnen het toelaatbare aantal (=30) parameters te blijven. Een eerste selectie is gemaakt op grond van een anova correlatietest van de KLA (kleur van het nagebakken product) met de overige parameters, waarmee één op één relaties met de KLA worden nagegaan. Hierbij wordt er van uitgegaan, dat als geen relatie bestaat tussen de bewuste parameter met de KLA, dit ook al in de anova test duidelijk is. Omgekeerd, als er wel een relatie bestaat zal deze naar voren moeten komen in de anova-test door wel enige correlatie te vertonen met de KLA.
In deze set, waarbij grofweg die parameters vallen die meer dan 10% positieve of negatieve correlatie vertonen met de KLA, vallen tevens alle parameters die in de literatuur vermeld staan en mogelijk in verband staan met de bakkleur. De geselecteerde parameters staan in tabel 4.

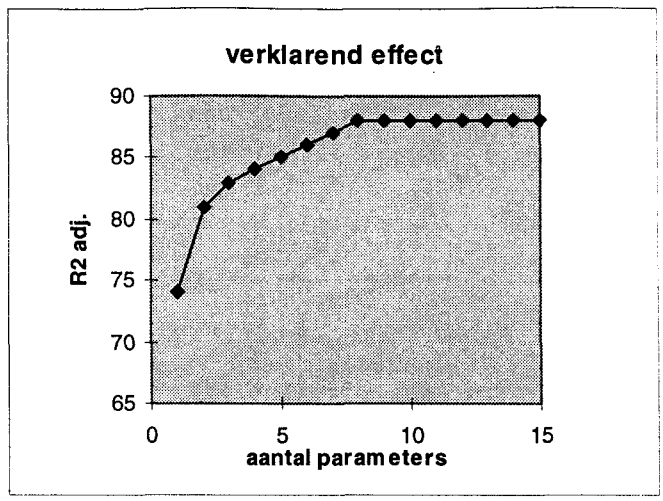
Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Tabel 4. Geselecteerde parameters door middel van de Anova-test. De bijbehorende correlaties zijn weergegeven als R^2. De code en omschrijving zijn zoals beschreven in het verslag "Methoden voor de karakterisering van de grondstof Aardappel".

code	omschrijving	R^2		code	omschrijving	R^2
101[1]	OWG	-0.40		114	glucose	0.83
104[2]	var. in knollen	-0.11		116	sucrose	0.27
105[13]	NMR	-0.22		118[2]	vit.C	0.20
107	zetmeel	-0.45		119[1]	aminoz.	0.29
109[1]	pH titr.	-0.38		119[2]	oplb.eiwit	-0.45
109[3]	pH midp.	-0.32		121	AIS	-0.30
110[1]	Na	0.23		208[2]	maat>70	0.12
110[2]	K	0.12		212	T100	0.60
110[5]	Fe	-0.12		213[1]	DSC dH	-0.43
110[6]	Cl	-0.27		213[2]	DSC onset	-0.27
112[1]	oxaalz.	-0.23		215[1]	chipskl.	-0.76
112[2]	citroenz.	-0.63		300	peroxidase	0.24
112[4]	fumaarz.	-0.22		301	PPO	-0.13
112[5]	pyrogl.z.	0.21		302	PE	0.20
113	KLI	0.86		303	protease	-0.17

Uitgaande van deze set is nagegaan met behulp van het statistisch programma Genstat, welke parameters, positief of negatief, een significant verband met de KLA vertonen. De meerwaarde van deze statistische test is dat de relaties nu niet meer alleen één op één met de KLA worden bepaald, maar ook de totale bijdrage met meerdere factoren tegelijk zichtbaar wordt.
Hieruit zijn 10 parameters naar voren gekomen als zijnde significant in de correlatie met elkaar tot de KLA. Indien meer parameters in de analyse worden betrokken leidt dit tot een vermindering van de gevonden correlatie.
De gevonden 10 significant bijdragende parameters hebben niet allen dezelfde bijdrage. De eerste twee parameters verklaren ca.82% van de correlatie en door het meenemen van meer parameters wordt de correlatie weliswaar significant hoger, maar het zijn toenames die in de orde van 1% of minder per gemeten parameter liggen (zie figuur 1). Op de Y-as staat nu de "R^2 adjusted" die in het algemeen iets lagere waarden geeft dan de R^2, maar wel een iets hogere statistische betrouwbaarheid heeft.
Het is duidelijk dat het maximaal haalbare 88% is, en dat de meeste winst zit in het gebruik van de eerste twee parameters (KLI en glucose).

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



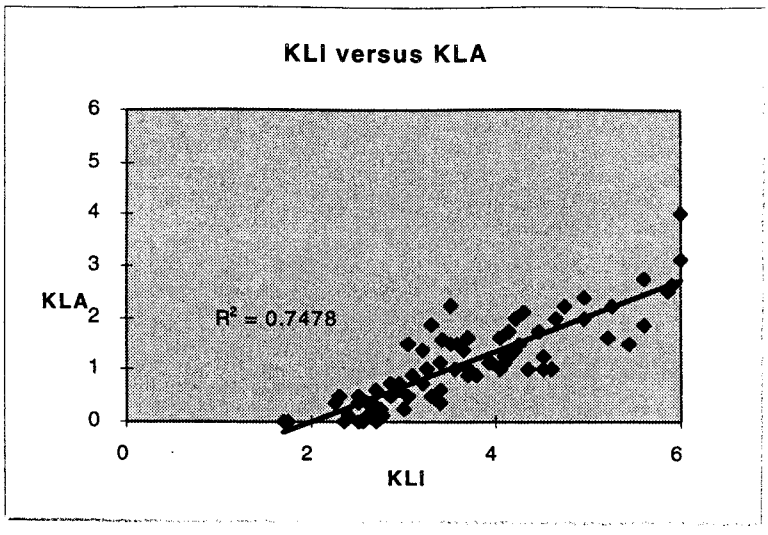
Figuur 1. Verklarend effect versus het aantal parameters.

Tabel 5. De geselecteerde parameters die in de statistische analyse zijn betrokken..

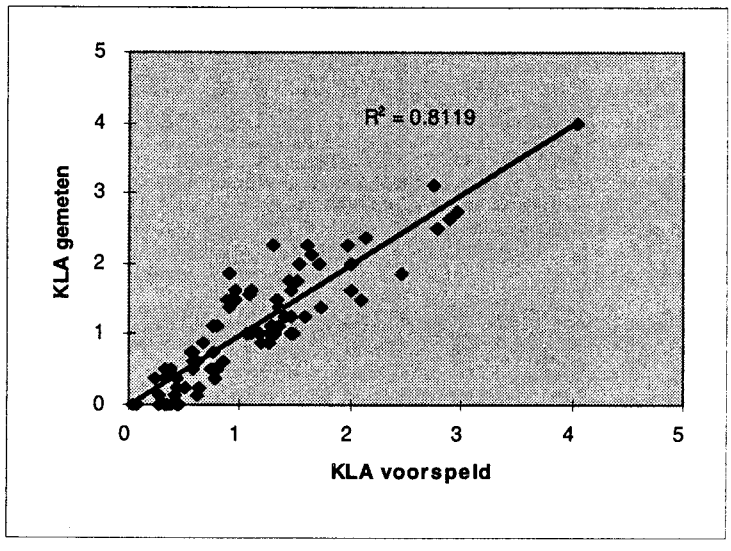
code	omschrijving		code	omschrijving
101[1]	OWG		213[1]	DSC dH
113	KLI		213[2]	DSC onset
114	glucose		215[1]	chipskl.
118[2]	vitam.C		300	peroxi-dase
119[2]	oplb. eiwit		301	PPO

De combinatie KLI en glucose, alsmede de combinatie KLI, glucose, oplosbaar eiwitgehalte en chipskleur zijn vervolgens nog getest op hun verklarend effect per jaar en op het effect van bewaard zijn of vers. Hierbij is geen statistisch verschil geconstateerd. Tevens is statistisch nagegaan in hoeverre een ras (en dan met name Bintje en Agria) kan afwijken van het totaal beeld dat verkregen is door de analyse op alle partijen. Hierbij bleek dat als de factor "ras" meegenomen werd het model er niet beter van werd, sterker nog, het werd zelfs, statistisch gesproken, wat slechter.
In de figuren 2 t/m 4 staan resp. de relatie KLI versus KLA (ruwe data), de relatie KLI versus KLA (geoptimaliseerd) en de relatie zoals die door het beste model met KLI/glucose versus KLA is berekend.

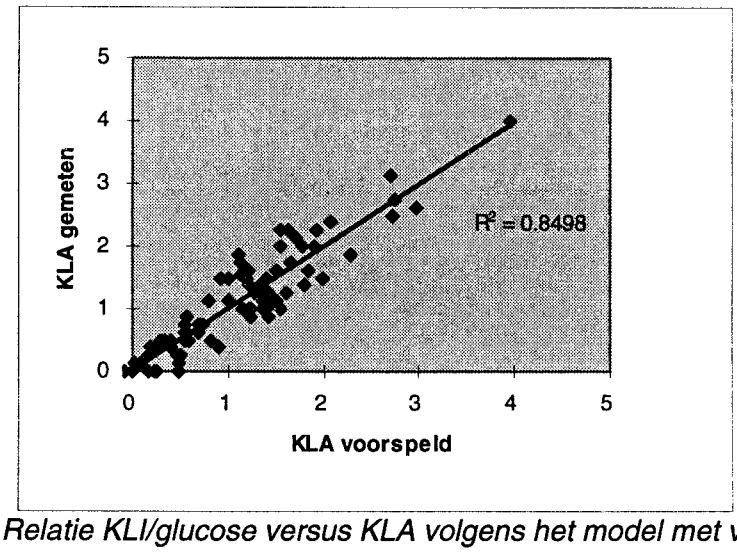
Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



Figuur 2. Relatie KLI versus KLA



Figuur 3. Relatie KLI versus KLA geoptimaliseerd voor twee parameters



Figuur 4. Relatie KLI/glucose versus KLA volgens het model met vier parameters.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

3.1.3 Conclusies

Het beste model met de twee parameters KLI en glucosegehalte geeft een goede verbetering ten opzichte van alleen KLI. Verdere verfijning met vier parameters geeft wel een statistische verbetering, maar voor de praktijk wellicht te weinig winst gezien de meetinspanning die er tegenover moet staan om vier parameters te meten.
Het eerste doel is nu een snelle, eenvoudige en robuuste glucosebepaling te vinden die vergelijkbare resultaten geeft als de laboratoriummethode zoals die in dit onderzoek zijn gebruikt.

3.2 Grauwverkleuring

3.2.1 Grauwverkleuring van aardappelen

Grauwverkleuring is een ernstig kwaliteitsgebrek van voorgebakken frites. De potentie van aardappel om grauw te verkleuren zit opgesloten in de aardappeleigenschappen. Grauwverkleuring van aardappelen wordt toegeschreven aan de reactie van Ijzer (Fe) met difenolen. In aardappel is het belangrijkste difenol chlorogeenzuur. Enzymatische reacties leveren ook een bijdrage aan de grauwverkleuring. In de betreffende enzymatische reactie worden fenolen in aanwezigheid van zuurstof omgezet tot polyfenolen en vervolgens tot melanoïden door het enzym polyfenoloxydase.
Teneinde de grauwverkleuring te voorkomen/remmen worden de geblancheerde frites in zuur natriumpyrofosfaat (SAPP) gedipt. De werking van pyrofosfaat berust op de complexering van het vrije ijzer, dat aanwezig is in de aardappelen: de reactie van het ijzer (Fe) met chlorogeenzuur wordt daarmee voorkomen. Het is niet duidelijk of pyrofosfaat een effect heeft op de enzymatische bijdrage. In de praktijk wordt PPi gedoseerd op basis van het criterium: 'afwezigheid van grauwverkleuring van de voorgebakken frites'. Dit houdt in dat PPi soms en mogelijk vaak overbodig wordt toegepast, wat leidt tot een onnodige milieubelasting met PPi.
Kennis van de grauwverkleuringspotentie van aardappelenpartijen zal bijdragen tot verlaging van het pyrofosfaatgebruik en daarmee milieubelasting
Voor optimale beheersing van de grauwverkleuring, zonderovermatig gebruik van pyrofosfaat, is het nodig dat:

1. de grauwverkleuringspotentie kan worden vastgesteld op basis van (bio)chemische karakterisering van de grondstof;
2. het effect van de procesomstandigheden op, gegeven de grondstof eigenschappen, op de grauwontwikkeling bekend is;
3. de effectiviteit van pyrofosfaat, kan worden vastgesteld op basis van (bio)chemische karakterisering van de grondstof;
4. methoden voor handen zijn de dosering van het pyrofosfaat adequaat te controleren (meten en regelen).

In dit verslag zijn de statistische resultaten weergegeven van de experimenten m.b.t. voorspellen van grauwverkleuring van voorgebakken frites op basis van de grondstofeigenschappen. (punt 1). De voorspelbaarheid van de pyrofosfaateffectiviteit was zeer laag: maximaal verklaarde variantie percentage was 36%, met 10 verklarende variabelen. Vandaar dat in dit verslag niet verder daarop zal worden ingegaan.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

De grondstof werd gekarakteriseerd op basis van een groot aantal, chemische, biochemische en fysische eigenschappen. De meetmethoden voor meten van de diverse eigenschappe zijn beschreven in rapport "Methoden voor de karakterisering van de grondstof aardappel"
Voor definitie en verduidelijking van de gebruikte statistische benaderingen c.q. termen wordt verwezen naar hoofdstuk 2 en bijlage 1.

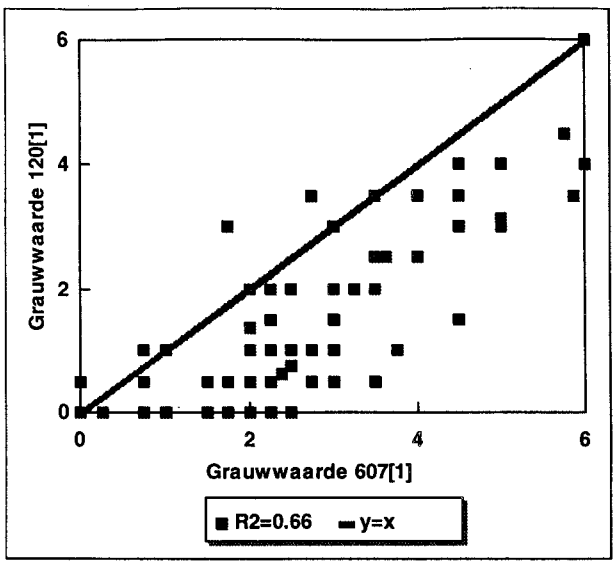
3.2.2 Grauwverkleuring en selectie van grondstofeigenschappen

Grauwverkleuring van de verwerkte partijen
Grauwverkleuring van de voorgebakken frites werd gemeten volgens methode 607[1] (zie Methoden-rapport) en volgens methode 120[1]. Methode 120[1] wordt gebruikt voor voorspelling van frites. In fig. 5 is het verband tussen beide methoden weergegeven voor 60 seconden voorbaktijd. De correlatie coëfficiënt voor dit verband r^2 is 0.66.

De grauwwaarde 607[1] blijkt over het algemeen hoger te zijn dan de grauwwaarde 120[1]. Dit is hoogstwaarschijnlijk een gevolg van het verschil in verwerkingsprocedure tussen de twee methoden. M.a.w. enige voorzichtigheid m.b.t. de absolute waarde van de grauwverkleuring is op zijn plaats, omdat de verwerkingsprocedure van invloed is op de waarde van de gemeten grauwverkleuring

Relevante componenten voor de grauwverkleuring
Op basis van theoretische kennis m.b.t. ontwikkeling van grauwverkleuring werden een aantal grondstofeigenschappen gekozen die mogelijk voorspellende waarde hebben. Deze componenten zijn tot een aantal van 30 aangevuld met eigenschappen die door de industrie aan de grondstof bepaald worden en een aantal chemische componenten, waarvan geen direct verband met de grauwverkleuring kan worden aangegeven (zie tabel 6). Het aantal van 30 is het maximum dat in de statistische analyse meegenomen kon worden. Met de zo verkregen 30 componenten, werd een lineaire regressie analyse uitgevoerd tussen enerzijds grauwverkleuring (607[1]) en anderzijds de diverse grondstofeigenschappen.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



Figuur 5: Verband tussen grauwwaarde van de frites verwerkt op de pilotlijn en de frites verwerkt middels de labmethode (120[1])

Tabel 6: Mogelijk relevante componenten gemeten in aardappelen m.b.t. grauwverkleuring bepaald met regressie analyse p= primaire eigenschap, s= secundaire (samengestelde) eigenschap, N= Niet-enzymatisch, E= Enzymatisch.

Nr.	Component	Methode	Verklaring
1	O.W.G (s)	101[1]	Maat voor drogestof, beïnvloedt de concentratie v/d componenten
2	variantieverdeling . s.g. binnen aardappels	104[2]	Industriële meting
3	NMR-vrij water	105[12]	Vrijwater kan de mobiliteit van componenten beïnvloeden
4	Zetmeegehalte	107	Industriële meting
5	Buffercapaciteit	109[1]	Kan de evenwichtsinstellingen van de Fe-Chl complexen beïnvloeden
6	pH Zuurgraad	109[1]	Kan de evenwichten en reactiesnelheden beïnvloeden
7	[Natrium] (p)	110[1]	Kan in competitie zijn met Fe bij complexvormingsreacties
8	[Ijzer] (p)	110[5]	Reageert met chlorogeenzuur tot 'grauwe componenten'
9	[fosfaat]	110[8]	Kan de complexvorming tussen ijzer en chlorogeenzuur beïnvloeden
10	[sulfaat]	110[9]	?
11	[Citroenzuur] (p)	112[2]	Complexeert ijzer en inhibeert PPO

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Nr.	Component	Methode	Verklaring
12	[Appelzuur] (p)	112[3]	Completeert ijzer?
13	[pyroglutamine zuur]	112[5]	?
14	frite skleurindex	113	Industriële meting, achtergrondkleur
15	[sucrose]	116	?
16	[Totaal fenolen] (p)	117[3]	Het substraat van PPO
17	[Chlorogeenzuur] (p)	118[1]	Belangrijkste chelaat van ijzer en substraat PPO
18	[Ascorbinezuur] (p)	118[2]	Inhibeert PPO en Completeert Fe
19	[Totaal zuur] (p)	118[3]	Beïnvloed zowel Fe-concentratie als PPO-activiteit
20	[Aminozuren] (p)	119[1]	?
21	[oplosbaar eiwit]	119[2]	Bezit fenolische groepen die kunnen bijdragen aan de kleur reacties
22	[Totaal eiwit] (p)	119[3]	Bevat tyrosine, het belangrijkste substraat van PPO in aardappelen.
23	Grauwmeting	120[1]	Voorspellend?
24	%AIS	121	?
25	% knollen > 70 mm	208[2]	Industriële meting
26	Afkookcijfer	210[1]	Industriëlemeting
27	Grauw verkleuring van gekookte aardappelen	210[2]	Voorspellende methode
28	Peroxydase-activiteit	300	?Gecorreleerd met PPO
29	PPO-activiteit (p)	301[1]	Katalyseert omzetting van de fenolen in zwart-bruine/blauwe verbindingen.
30	Protease-activiteit	301[2]	Continueert omzetting van fenolen , ook nadat geblancheerd is.

3.2.3 Statistische selectie van relevante componenten en modelvorming

Op basis van 'all-possible-subsets' regressies met de 30 geselecteerde grondstofeigenschappen, werden vier grondstofeigenschappen uitgekozen om een voorspellend model mee te maken. Deze zijn gegeven in tabel 7. Arbitrair is gekozen voor vier grondstofeigenschappen.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Verder blijkt dat:
- de gekozen grondstofeigenschappen een lineair verband vertonen met de grauwverkleuring;
- er geen significante interacties bestaan tussen de componenten.

Dit betekend dat, indien de betreffende grondstofeigenschap b.v. twee keer zo groot wordt, de bijdrage van deze grondstofeigenschap aan de grauwverkleuring twee keer zo groot (bij positief verband) of klein (bij negatief verband) wordt.

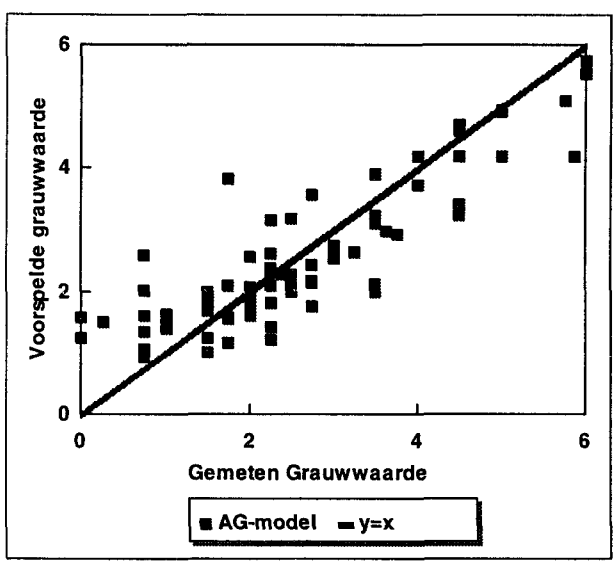
3.2.4 Algemeen Grauw-model

Tabel 7: Algemeen Grauw model voor voorspelling van grauwverkleuring.

	Parameter	Waarde	Standard Error	Verhouding (t-waarde)
C	Constant	-1.8	0.61	3
110[5]	Ijzer	386	101	3.8
118[1]	Chlorogeenzuur	10	4.2	2.4
118[3]	Totaal zuur	0.25	0.096	2.6
120[1]	Labmethode	0.68	0.07	9.7
	R^2adj	75.7%	S.E. Obs.	0.75

In formule vorm ziet het AG-Model er als volgt uit:

Grauw = -1.8•C + 386•110[5] + 10•118[1] + 0.25•118[3] + 0.68•120[1]



Figuur 6: Statistisch model van de relatie tussen een viertal grondstofeigenschappen en de grauwwaarde 607[1]

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

De verhouding van de waarde van de parameter en de standaard error voor de parameter (zie kolom 4) is een maat voor het belang van die component in het model. Te zien is dat methode 120[1] de grootste bijdrage geeft en vervolgens ijzer. Chlorogeenzuur en Totaal zuur hebben een vergelijkbare bijdrage.
De voorspellende kracht van het model is weergegeven m.b.v. R^2adj en S.E. Obs. In vergelijking met het Model 120[1] (zie par. 3.2.4) geeft dit model een significant betere voorspelling. Het resultaat van de voorspelling is weergegeven in figuur 6. Te zien is de spreiding van de punten nogal groot is. Dit geeft aan dat de standaard fout (standard error) van het model relatief groot is. Aan de ligging van de punten boven de lijn y=x, kan geconcludeerd worden dat de grauwwaarde van de aardappelen met een lage grauwwaarde (vooral Agria) wordt overschat. De grauwwaarde van de aardappelen met een hoge grauwwaarde (vooral Saturna) worden daarentegen onderschat.

3.2.5 Effect van bewaring

Teneinde na te gaan of er een significant effect is van de leeftijd (vers t.o.v. bewaard) van de aardappelen op de voorspelbaarheid m.b.v. de vier genoemde grondstofeigenschappen is deze als factor in de analyse betrokken. Het resultaat is weergegeven in tabel 8.

Tabel 8: Leeftijd-model voor voorspelling van grauwverkleuring van verse en bewaarde aardappelen

		Vers	Bewaard
	Parameter	Waarde	Waarde
C	Constant	-2.9	-0.4
110[5]	Ijzer	581	231
118[1]	Chlorogeenzuur	11.7	10.2
118[3]	Totaal zuur	0.27	0.17
120[1]	Labmethode	0.72	0.62
	R^2adj	76.6%	S.E. Obs. 0.73

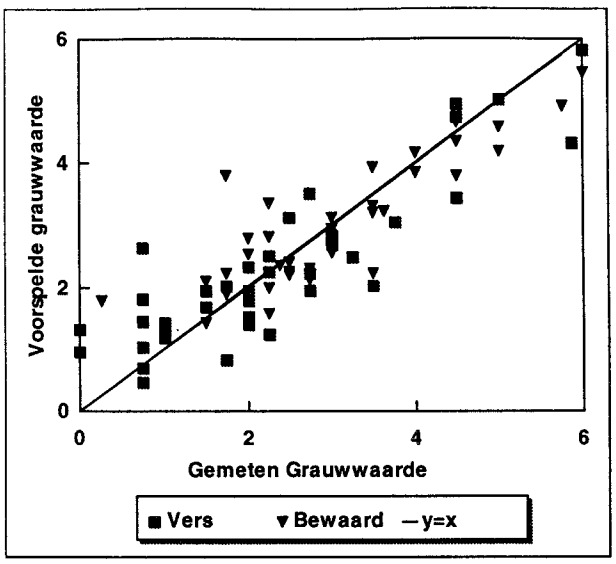
Uit het grote verschil tussen de parameterwaarden van ijzer (150%) en Totaal zuur (60%) kan geconcludeerd worden dat deze componenten bij verse aardappelen een sterkere rol spelen dan bij bewaarde aardappels. Vergelijking van dit model met het AG-model geeft aan dat er een kleine maar significante verbetering verkregen wordt door meenemen van de factor leeftijd.

Evenals in het vorige model is er wederom een relatief grote spreiding te zien en ook, zij het iets minder, een overschatting en onderschatting van het lage respectievelijk hoge grauwwaarden.

3.2.6 Effect van oogstseizoen

Teneinde na te gaan of er een significant effect is van het oogstseizoen op de voorspelbaarheid m.b.v. de vier genoemdegrondstofeigenschappen is het oogstjaar ('93, '94, '95) als factor in de analyse betrokken. Het resultaat is weergegeven in tabel 9.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



Figuur 7: Statistisch model voor de relatie tussen een viertal grondstofeigenschappen en de grauwwaarde 607[1], waarbij de leeftijd als factor is meegenomen.

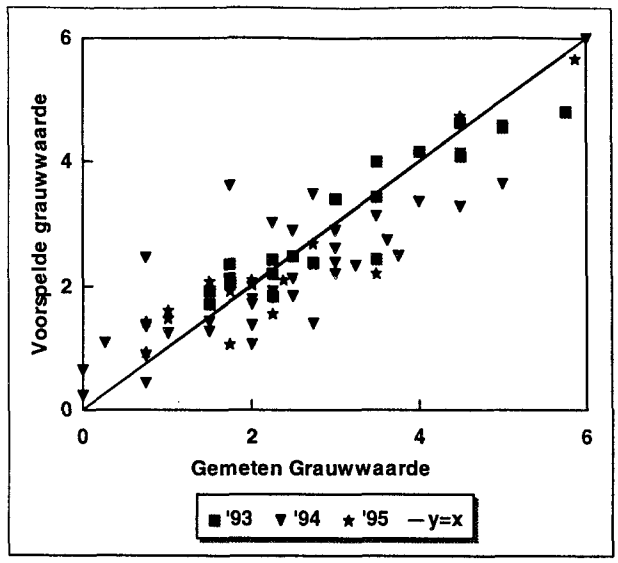
Tabel 9: Algemeen Grauw model voor voorspelling van grauwverkleuring

		1993	1994	1995
	Parameter	Waarde	Waarde	Waarde
C	Constant	0.47	-2.7	1.2
110[5]	Ijzer	181	757	438
118[1]	Chlorogeenzuur	10.8	11.6	8.4
118[3]	Totaal zuur	0.091	0.20	-0.25
120[1]	Labmethode	0.58	0.78	0.92
	R^2adj	81.1%	S.E. Obs.	0.66

Ook in dit geval blijkt dat de bijdrage van vooral ijzer en totaal zuur sterk afhankelijk is van het oogstjaar. Vergelijking van dit model met de voorgaande modellen geeft aan dat zowel t.o.v. het AG-model als het leeftijd-model een significante verbetering verkregen wordt door meenemen van de factor oogstjaar.

Het algemene beeld (zie figuur 8) is dat er ook in dit model sprake is van overschatting bij lage waarden en onderschatting bij hoge waarden. Echter, de mate van overschatting en onderschatting is echter aanzienlijk kleiner dan bij het AG-model.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



Figuur 8: Statistisch model van de relatie tussen een viertal grondstofeigenschappen en grauwwaarde 607[1], waarbij de factor oogstjaar is meegenomen.

3.2.7 Ras: Agria versus Bintje

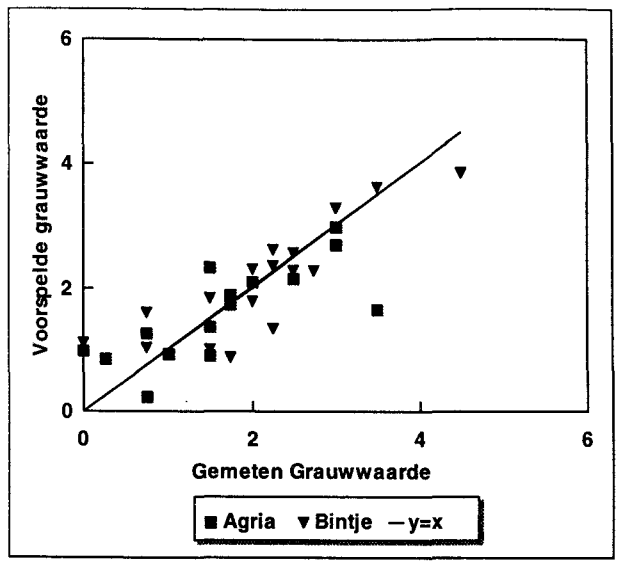
Teneinde na te gaan of m.b.v. deze grondstofeigenschappen ook een dergelijk model gemaakt kan worden is de analyse ook uitgevoerd voor de 15 partijen Agria en de 17 partijen Bintje. Het resultaat is weergegeven in tabel 10.

Tabel 10: Ras-model model voor voorspelling van grauwverkleuring van Agria en Bintje

		Agria+Bintje	Agria (15 Partijen)	Bintje (17 Partijen)
	Parameter	Waarde	Waarde	Waarde
C	Constant	-2.8	-7.8	-2.4
110[5]	Ijzer	661	705	744
118[1]	Chlorogeenzuur	4.8	24.8	12.3
118[3]	Totaal zuur	0.24	0.92	0.125
120[1]	Labmethode	0.80	0.6	0.61
	R^2adj	?	S.e. obs.	?

R^2adj en S.e. obs. is in dit model niet gegeven omdat het niet bruikbaar is voor vergelijking met de overige modellen. De reden hiervan is dat de twee geselecteerde rassen eigenschappen hebben die niet de gehele range aan waarden hebben: b.v. de range van Agria voor grauw is 0-3,5 en de range voor Bintje is 1-4,5. Desalniettemin in figuur 9 het resultaat van het bijbehorende model weergegeven.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



Figuur 9: Statistisch model van de relatie tussen een viertal grondstofeigenschappen en de grauwwaarde 607[1] voor Agria en Bintje.

Te zien is dat dit ook dit model een relatief grote spreiding geeft. De spreiding is ongeveer hetzelfde voor beide rassen. De overschatting en onderschatting voor lage en hoge grauwwaarden is niet zichtbaar.

3.2.8 Conclusie

Van het groot aantal gemeten grondstofeigenschappen bleken drie een significante bijdrage te leveren aan de voorspelling: Ijzer, Chlorogeenzuur, totaal zuur. De beste voorspelling van de grauwwaarde, gemeten aan de frites verwerkt op de pilotlijn, wordt verkregen indien de grauwwaarde van de frites verwerkt op labschaal (methode 120[1]) in het model wordt betrokken. Deze laatste methode herbergt gedeeltelijk al in zich de bijdrage van de drie chemische componenten en tevens de mogelijke bijdrage van de enzymatische reacties.

De voorspellende kracht van het algemene model, gebaseerd op alle partijen, is echter nog te laag vanwege de relatief lage verklarende waarde R^2adj en de relatief hoge standaard fout.

Inbreng van de factor leeftijd (vers of bewaard) of oogstjaar geeft weliswaar een significante verbetering, maar leidt niet tot een bruikbaar model.

3.3 Textuur

3.3.1 Inleiding

Het eenduidig vaststellen van de textuur van frites blijkt geen eenvoudige aangelegenheid te zijn, wat ook al gebleken was uit textuurbeoordelingen in de praktijk. Eveneens is duidelijk dat grote verschillen in nauwkeurigheid bestaan voor de verschillende textuuraspecten. Zo stellen de productexperts bijvoorbeeld redelijk eenduidig een waarde vast voor TXA (nat/droog) en TXB (hard/slap). Het analytisch sensorisch panel doet hetzelfde voor bijvoorbeeld stevigheid en krokantheid. Daarentegen wordt aan het aspect taaiheid van frites een minder eenduidige waarde

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

gegeven. In dit deelonderzoek naar statistische relaties zal verslag gedaan worden van de gevonden correlaties tussen verschillende textuuraspecten van het product met één of meerdere grondstofparameters.

3.3.2 Methoden

De gebruikte analysemethoden voor de metingen van de verschillende grondstofkenmerken zijn samengevat in het verslag "Methoden voor de karakterisering van de grondstof aardappel" (versie januari 1996). Textuureigenschappen zijn in dit onderzoeksproject beoordeeld door enerzijds een analytisch sensorisch panel (20-25 personen) en anderzijds twee productexperts. Beiden gebruiken eigen termen voor de verschillende textuuraspecten. Voor de gebruikte statistische methoden wordt verwezen naar het algemene hoofdstuk in dit verslag. De modellen die hier besproken worden, zijn tot stand gekomen op basis van de onderzoeksresultaten t/m het vijfde verwerkingstijdstip. Daarbij zijn alleen de resultaten van frituurtijden van 60 seconden meegenomen.

3.3.3 Resultaten

Het kiezen van grondstofparameters t.b.v. voorspellende modellen

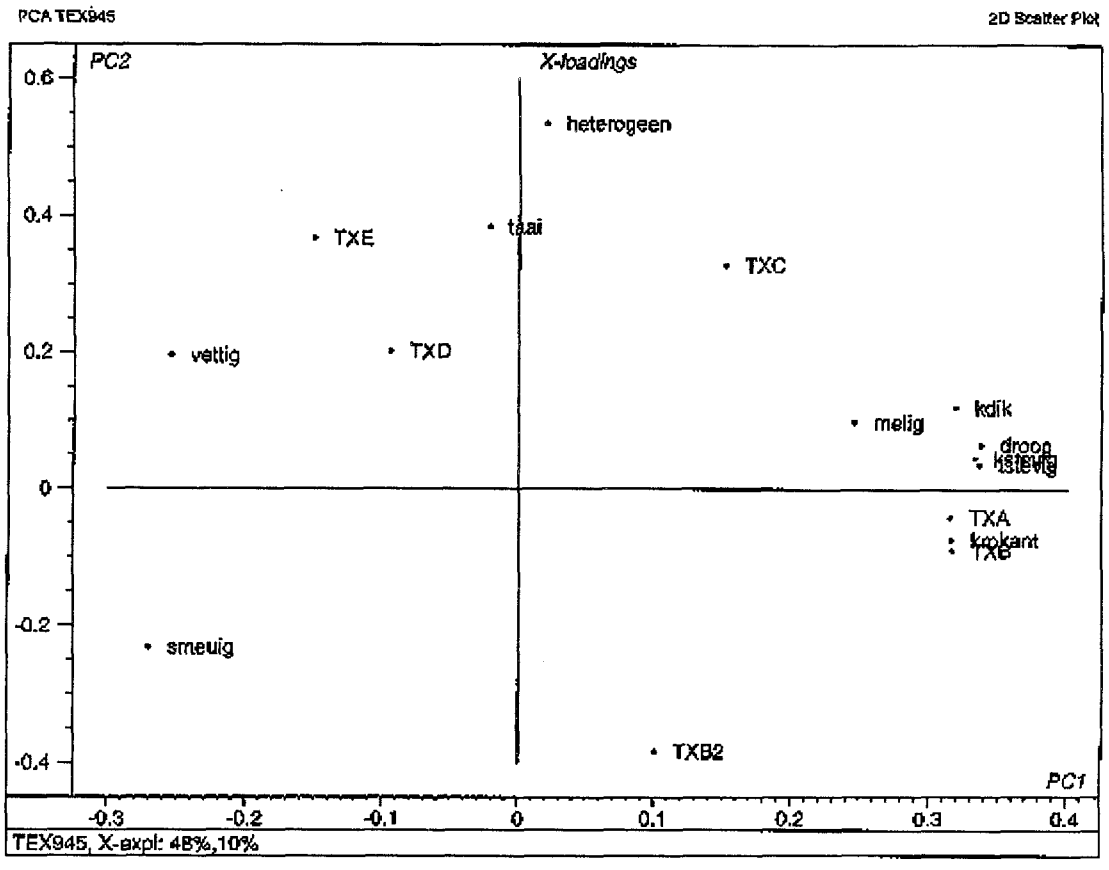
Van de meer dan 100 gemeten grondstofparameters gedurende het totale onderzoek zijn 57 parameters ieder verwerkingstijdstip tot en met het vijfde gemeten. Uit deze 57 grondstofparameters is een selectie van 30 stuks gemaakt om zo binnen het maximaal statistisch (programma) toetsbare aantal parameters te blijven. Deze 30 parameters (zie tabel 11) omvatten alle methoden zoals gebruikt in de praktijk (onderwatergewicht, friteskleurindex en maatsortering) alsmede methoden die relevant geacht worden ten aanzien van het ontstaan van de textuur van het uiteindelijke product. Hierbij is van parameters die onderling een hoge correlatie vertoonden - zoals bv. het onderwatergewicht en het drogestofgehalte - steeds maar één parameter opgenomen.

Tabel 11: Selectie van 30 grondstofparameters t.b.v. de textuuraspecten. Genoemde codes en omschrijvingen zijn terug te vinden in het verslag "Methoden voor de karakterisering van de grondstof aardappel".

Code	Omschrijving	Code	Omschrijving	Code	Omschrijving
101[1]	OWG	110[2]	kalium	119[3]	totaal N
103[2]	var. sg. tussen aardappelen	110[3]	calcium	121	% AIS
103[4]	%knollen 1.11	110[6]	chloride	201[1]	celgrootte
104[2]	var. sg. binnen aardappelen	110[9]	sulfaat	208[2]	maatsortering %knollen > 70
104[4]	%knollen 1.115	112[1]	oxaalzuur	210[1]	afkookcijfer
105[12]	NMR	112[2]	citroenzuur	212[1]	T100
107	zetmeel	112[3]	appelzuur	213[1]	DSC dH
109[2]	pH bufferconc.	112[4]	fumaarzuur	300[1]	POD
109[4]	pH zuurgraad	112[5]	pyroglutamine-zuur	302	PE
110[1]	natrium	113	friteskleurindex	303	protease

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Uit alle textuuraspecten zijn vijf aspecten gekozen voor nader statistisch onderzoek. Deze vijf aspecten zijn gekozen op basis van een Principle Component Analyse. Van aspecten die onderling een hoge correlatie vertoonden - zoals bv. TXA (nat/droog) en TXB (slap/hard) - is steeds maar één aspect opgenomen. De vijf aspecten liggen verspreid over het PCA-plaatje en geven derhalve een goed beeld van de totaal waargenomen textuur van frites.



Figuur 10 PCA-analyse van textuureigenschappen

In tabel 12 wordt een overzicht gegeven van alle textuuraspecten die in het project "Karakterisering grondstof aardappel en voorspelling eigenschappen bij verwerking" bepaald zijn. In dit verslag zullen de volgende vijf aspecten besproken worden: **TXA** (nat/droog) en **TXC** (taai) zoals beoordeeld door productexperts en **krokante korst**, **smeuïgheid** en **heterogeniteit** zoals beoordeeld door een analytisch sensorisch panel.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

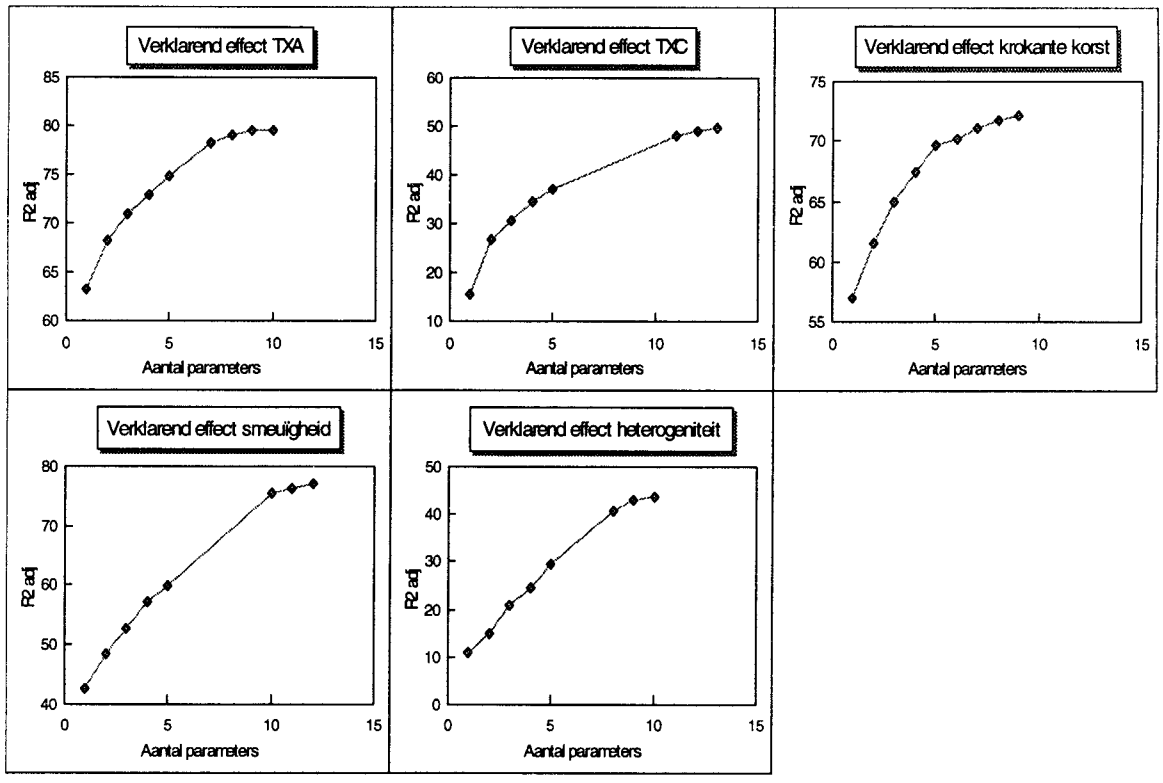
Tabel 12: Verklaring van alle textuuraspecten zoals bepaald door productexperts en het analytisch sensorisch panel in het project "Karakterisering grondstof aardappel en voorspelling eigenschappen bij verwerking".

Code	Textuuraspect	Verklaring	Uitvoering door
603[1]	BLA	hoeveelheid blaasjes aanwezig op frites op een schaal van 0 (geen blaasjes) tot 2 (blaasjes)	productexperts
603[2]	***TXA (nat/droog)***	***textuur binnenkant: mate van nat of droog zijn van frites op een schaal van -2 (nat) tot 2 (droog)***	***productexperts***
603[3]	TXB (slap/hard)	textuur buitenkant: mate van slap of hard zijn van frites op een schaal van -2 (slap) tot 2 (hard)	productexperts
603[4]	TXB2 (hol)	holheid van frites op een schaal van 0 (niet hol) tot 2 (hol)	productexperts
603[5]	***TXC (taai)***	***de mate van taaiheid van frites op een schaal van 0 (niet taai) tot 2 (taai)***	***productexperts***
603[6]	TXD (gaar)	de mate van gaarheid van frites op een schaal van 0 (gaar) tot 2 (niet gaar)	productexperts
603[7]	SMK (smaak)	de mate van off-flavours in frites op een schaal van 0 (geen off-flavour) tot 5 (sterke off-flavour)	productexperts
603[8]	TXE (heterogeniteit)	de heterogeniteit van de textuur op een schaal van 0 (niet heterogeen) tot 2 (heterogeen)	productexperts
603[9]	FE (feathering)	de mate van feathering op frites op een schaal van 0 (geen feathering) tot 10 (sterke feathering)	productexperts
604[11]	***krokante korst***	***korst van frites gemakkelijk breekbaar (knappend)***	***analytisch sensorisch panel***
604[12]	taaie korst	korst kauwbaar en vervormbaar, breekt niet direct in stukjes	analytisch sensorisch panel
604[13]	stevige korst	weerstand korst bij initiële beet	analytisch sensorisch panel
604[31]	dikke korst	dikke korst, stukjes korst minder schilferig	analytisch sensorisch panel
604[32]	meligheid	fijne fritesstructuur, meer speeksel nodig	analytisch sensorisch panel
604[33]	droogheid	droge fritessubstantie, meer speeksel nodig	analytisch sensorisch panel
604[34]	vettigheid	vettige flim in mond bij eten van frites	analytisch sensorisch panel
604[35]	***smeuïgheid***	***gladde, vaak fijne fritesstructuur, glijdt gemakkelijk en plakt niet***	***analytisch sensorisch panel***
604[36]	***heterogeniteit***	***verschil in aspecten tussen frites binnen één monster***	***analytisch sensorisch panel***
604[37]	stevigheid	weerstand bij kauwen van frites	analytisch sensorisch panel

NB: De vijf aspecten - gekozen voor het statistisch onderzoek - zijn vet en cursief weergegeven.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Uitgaande van de set van 30 grondstofparameters (zie tabel 11) is nagegaan welke van deze parameters een significant verband met één van de vijf textuuraspecten vertoont. De hiervoor gebruikte statistische analyses zijn uitgevoerd met behulp van Genstat v5.32. Het doel van deze statistische analyses is het vinden van geschikte modellen voor de vijf genoemde textuuraspecten, waarmee de textuur vanuit de grondstof voorspeld kan worden. In zo'n model wordt gebruik gemaakt van een aantal grondstofparameters. Het is eenvoudig te beredeneren dat een voorspelling van de textuur vanuit een groot aantal grondstofparameters nauwkeuriger zal zijn dan een voorspelling vanuit 1 of 2 grondstofparameters. Om hierin enig inzicht te verkrijgen zijn "all-possible-subset regressies" uitgevoerd, zodat een overzicht wordt verkregen van de beste modellen met 1, 2, 3 of meer verklarende parameters. Figuur 11 geeft een indruk van het effect van het aantal grondstofparameters op het verklarend effect (R^2_{adj}) van de verschillende textuuraspecten.



Figuur 11 Effect van het aantal grondstofparameters in het voorspellend model op het verklarend effect (R^2_{adj}) van de verschillende textuuraspecten (TXA, TXC, krokante korst, smeuïgheid en heterogeniteit).

Uit figuur 11 blijkt dat het verklarend effect van de verschillende textuuraspecten beperkt is (vooral ten aanzien van TXC en heterogeniteit), ook wanneer een model van tien parameters wordt toegepast. Het lijkt dus niet eenvoudig om vanuit (een beperkt aantal) grondstofparameters de textuur van frites te voorspellen. Ten behoeve van het statistisch onderzoek met betrekking tot de textuur van frites is gekozen voor voorspellende modellen met vier grondstofparameters. De keuze van vier parameters is een compromis tussen de praktische realiseerbaarheid (het bepalen van vier versus bv. tien grondstofparameters) en het verklarend effect ervan.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

3.3.4 Voorspellende modellen

TXA (nat/droog)

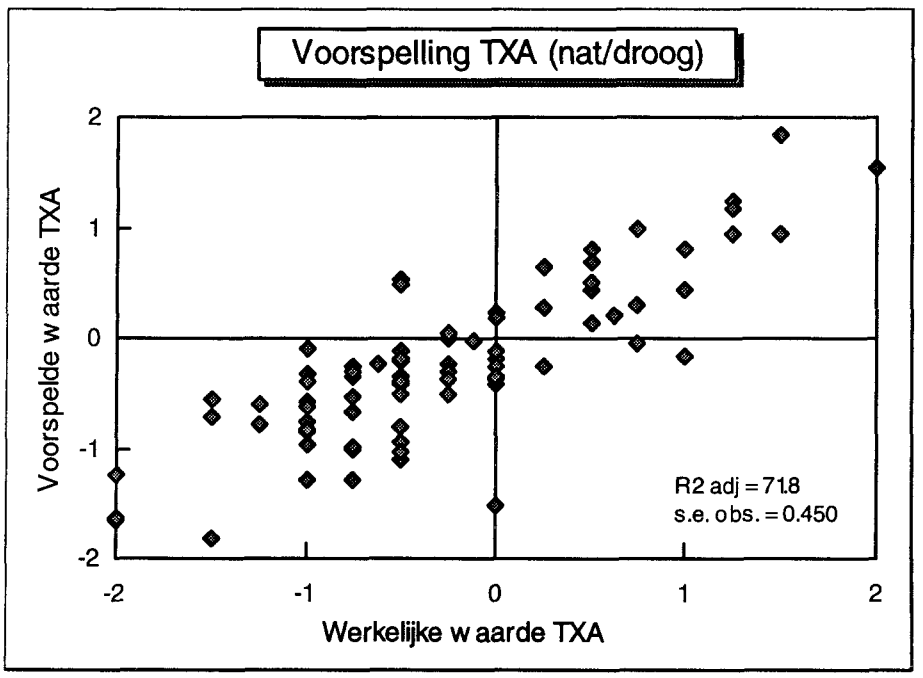
Op basis van een "all-possible-subset regressie" is voor TXA (nat/droog) gekozen voor de volgende vier grondstofparameters:

- 101[1] onderwatergewicht (OWG);
- 110[2] kaliumgehalte;
- 110[6] chloridegehalte;
- 119[3] totaal stikstofgehalte (N).

Met deze vier grondstofparameters is een regressie-analyse op TXA uitgevoerd. Het meenemen van interacties en/of kwadratische termen leidde hierbij niet tot een verbetering. Het uiteindelijke model wordt weergegeven door de volgende vergelijking:

TXA = -7.6 + 0.015*[101[1]] - 0.12*[110[2]] + 0.73*[110[6]] + 0.50*[119[3]]

In figuur 12 wordt de voorspelling van TXA met het ontwikkelde model weergegeven. Het percentage verklaarde variantie (R^2_{adj}) van dit model is 71,8%.



Figuur 12: Voorspelling van TXA (nat/droog) op basis van een vier-parameter-model

Tevens is gekeken naar het effect van een aantal factoren op de modelparameters, zijnde leeftijd (vers cq. bewaard), oogstjaar (1993 cq. 1994 cq. 1995) en ras (Agria cq. Bintje). Het meenemen van de factor leeftijd in het model leidt tot een significante, maar kleine verbetering van het model. Het percentage verklaarde variantie (R^2_{adj}) wordt in dit geval 74,2% en s.e. obs. wordt 0,430. Vooral het totaal stikstofgehalte, 119[3], lijkt bij verse aardappels een andere rol te spelen dan bij bewaarde aardappelen. Het meenemen van de factor oogstjaar geeft eveneens een significante verbetering van het model en verhoogt het percentage verklaarde variantie naar 74,4% en verlaagt de s.e.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

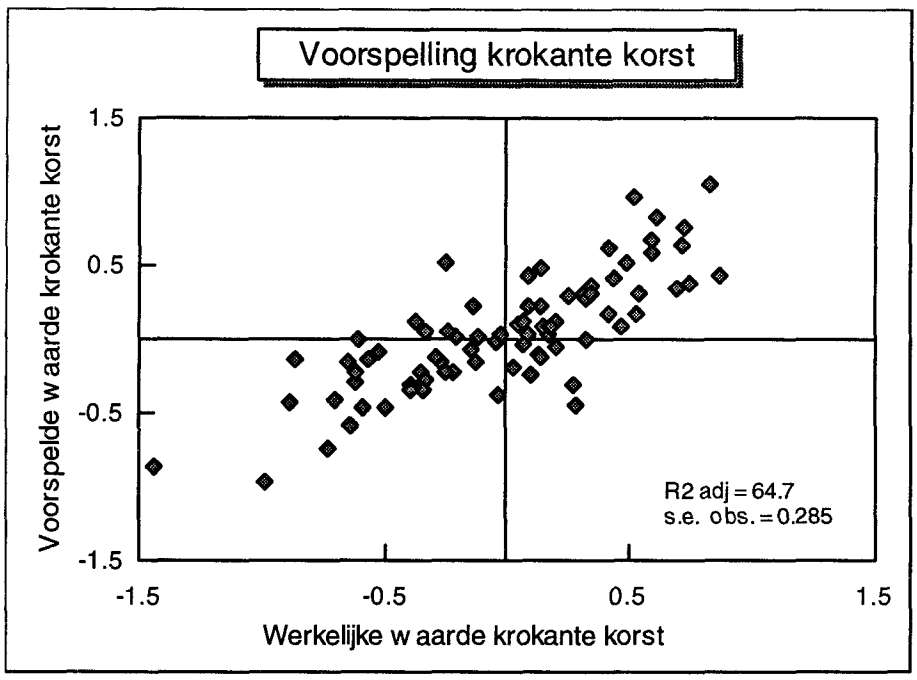
obs. naar 0,428. Vooral de invloed van kalium varieert per jaar. De verschillen tussen de twee rassen zijn niet significant, wat betekent dat het meenemen van de factor ras geen verbetering van het model voor TXA geeft. Overigens dient opgemerkt te worden dat de uitsplitsing naar de rassen Agria en Bintje relatief kleine datasets oplevert, waaruit niet zonder meer conclusies getrokken mogen worden met betrekking tot significantie.

Krokante korst
Op basis van een "all-possible-subset regressie" is voor krokante korst gekozen voor de volgende vier grondstofparameters:
- 101[1] onderwatergewicht (OWG);
- 104[2] variantie verdeling soortelijk gewicht tussen aardappels;
- 109[2] pH bufferconcentratie;
- 113 frituskleurindex.

Met deze vier grondstofparameters is een regressie-analyse op krokante korst uitgevoerd. Het meenemen van interacties en/of kwadratische termen leidde hierbij niet tot een verbetering. Het uiteindelijke model wordt weergegeven door de volgende vergelijking:

Krokante korst = -2.7 + 0.008*[101[1]] + 445*[104[2]] - 0.031*[109[2]] - 0.14*[113]

In figuur 13 wordt de voorspelling van krokante korst met het ontwikkelde model weergegeven. Het percentage verklaarde variantie (R^2_{adj}) van dit model is 64,7%.



Figuur 13: Voorspelling van krokante korst op basis van een vier-parameter-model.

Ook ten aanzien van krokante korst is gekeken naar het effect van een aantal factoren op de modelparameters, zijnde leeftijd (vers cq. bewaard), oogstjaar (1993 cq. 1994 cq.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

1995) en ras (Agria cq. Bintje). Een uitsplitsing per leeftijd levert geen significante leeftijdseffecten op het model op. Het meenemen van de factor oogstjaar geeft eveneens geen significante verbetering van het model. Tenslotte heeft ook een uitsplitsing per ras geen invloed op de verklarende waarde van het model.

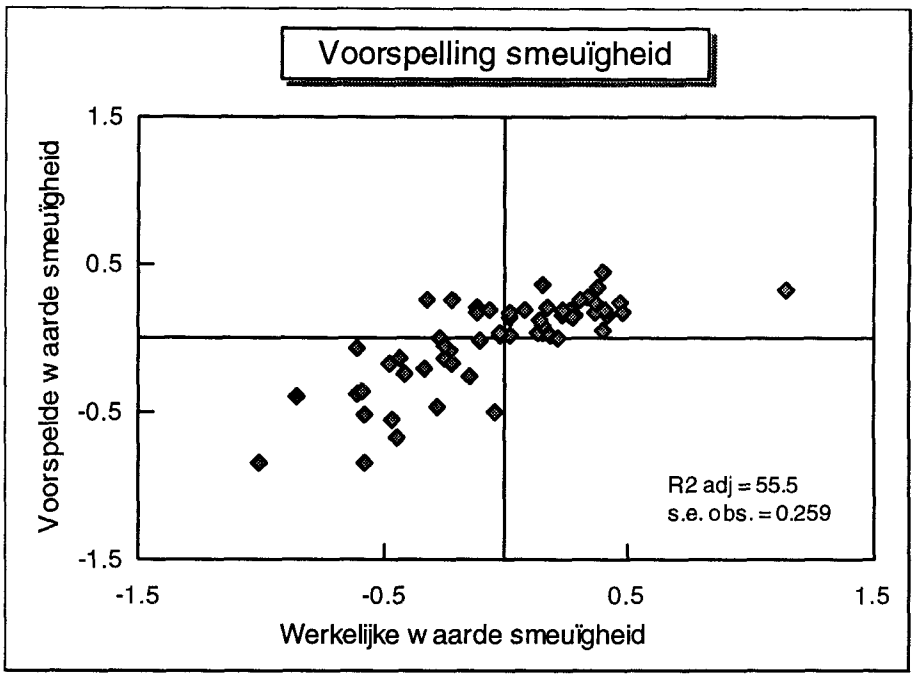
Smeuïgheid
Op basis van een all-possible-subset regressie is voor smeuïgheid gekozen voor de volgende vier grondstofparameters:
- 101[1] onderwatergewicht (OWG);
- 104[2] variantie verdeling soortelijk gewicht tussen aardappels;
- 110[6] chloridegehalte;
- 300 peroxidase activiteit (POD).

Opgemerkt dient te worden dat het textuuraspect smeuïgheid pas vanaf de tweede ronde is gemeten.

Met deze vier grondstofparameters is een regressie-analyse op smeuïgheid uitgevoerd. Het meenemen van interacties en/of kwadratische termen leidde hierbij niet tot een verbetering. Het uiteindelijke model wordt weergegeven door de volgende vergelijking:

Smeuïgheid = 2.9 - 0.0045*[101[1]] - 442*[104[2]] - 0.38*[110[6]] - 39*[300]

In figuur 14 wordt de voorspelling van smeuïgheid met het ontwikkelde model weergegeven. Het percentage verklaarde variantie (R^2_{adj}) van dit model is 55,5%.



Figuur 14: Voorspelling van smeuïgheid op basis van een vier-parameter-model.

Gekeken is naar het effect van een aantal factoren op de modelparameters, zijnde leeftijd (vers cq. bewaard), oogstjaar (1993 cq. 1994 cq. 1995) en ras (Agria cq. Bintje).

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Een onderscheid per leeftijd geeft een significante en grote verbetering van het model. Het percentage verklarende variantie (R^2_{adj}) wordt in dat geval namelijk verhoogd naar 63,3% en de s.e. obs. wordt verlaagd naar 0,23. De invloed van zowel het onderwatergewicht als de variantie verdeling soortelijk gewicht tussen aardappelen als de peroxidase-activiteit is voor verse aardappelen anders dan voor bewaarde aardapelen. Het meenemen van de factor oogstjaar geeft geen verbetering van het model. Eveneens bestaat geen significante rasinvloed op het model.

TXC (taai)
Op basis van een all-possible-subset regressie is voor TXC (taai) gekozen voor de volgende vier grondstofparameters:
- 105[12] NMR vers materiaal;
- 107 zetmeelgehalte;
- 110[6] chloridegehalte;
- 112[1] oxaalzuurgehalte.

Met deze vier grondstofparameters is een regressie-analyse op TXC uitgevoerd. Het meenemen van de kwadratische termen voor 107 en 112[1] leidde tot een verhoging van het percentage verklarende variantie (R^2_{adj}) van 35% tot 42%. Dit percentage blijft echter aan de lage kant. Het meenemen van de factoren leeftijd, oogstjaar en ras levert eveneens geen verbetering van het model. Verdere details van het uiteindelijke model worden hier niet gegeven.

Heterogeniteit
Op basis van een all-possible-subset regressie is voor heterogeniteit gekozen voor de volgende vier grondstofparameters:
- 107 zetmeelgehalte;
- 112[2] citroenzuurgehalte;
- 208[2] % knollen met maat > 70;
- 303 proteasegehalte.

Opgemerkt dient te worden dat het textuuraspect heterogeniteit pas vanaf de tweede ronde is gemeten.

Met deze vier grondstofparameters is een regressie-analyse op heterogeniteit uitgevoerd. Het meenemen van interacties en/of kwadratische termen leidde hierbij tot een verbetering van het model. Als de kwadratische term voor 107 en de interactie 107*303 worden meegenomen, gaat het percentage verklarende variantie (R^2_{adj}) van 24% naar 37%. Ook hier blijft echter het percentage verklarende variantie voor het model voor heterogeniteit aan de lage kant. Het meenemen van de factoren leeftijd en oogstjaar levert geen significante verbetering van het model op. Een uitsplitsing per ras verhoogt de R^2_{adj} van 32% naar 39%, maar het gaat hier wel om een zeer kleine dataset. Verdere details van het uiteindelijke model worden hier niet gegeven.

3.3.5 Onderwatergewicht versus drogestofgehalte
Het drogestofgehalte en het onderwatergewicht vertonen onderling een hoge correlatie. Zoals eerder aangegeven is in de set van 30 parameters (zie tabel 11) alleen het onderwatergewicht opgenomen, omdat dit in de praktijk gemeten wordt. Geanalyseerd is in hoeverre de voorspelling voor het textuuraspect TXA en de optimale meetmethoden wijzigen indien het drogestofgehalte zou worden bepaald in plaats van

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

het onderwatergewicht. Dezelfde statistische verwerkingen zijn uitgevoerd met een op één na identieke set van 30 parameters: het onderwatergewicht is vervangen door het drogestofgehalte. Dit leverde voor het 4-parameter-model voor TXA een hoger verklarend effect (R^2_{adj}) op, nl. 77,1% tegenover 71,8% in geval van onderwatergewicht. Eveneens werden andere mede verklarende variabelen gevonden (onderwatergewicht - kaliumgehalte - chloridegehalte - totaal stikstofgehalte tegenover drogestofgehalte - kaliumgehalte - fumaarzuurgehalte - celgrootte). Het drogestofgehalte heeft dus andere mede verklarende parameters dan het onderwatergewicht. Het gebruik van het drogestofgehalte levert weliswaar een hoger verklarend effect (in ieder geval ten aanzien van TXA), praktisch betekent het echter meer werk, daar de bepaling van het drogestofgehalte tijdrovender is dan de bepaling van het onderwatergewicht. Een afweging dient dus gemaakt te worden tussen een iets beter voorspellend model voor TXA (5% meer) en een meer tijdrovende grondstofbepaling. Op dit moment is gekozen voor modellen uitgaande van het onderwatergewicht.

3.3.6 Conclusie

Van de vijf onderzochte textuuraspecten zijn TXA (nat/droog), krokante korst en smeuïgheid het best voorspelbaar. Uit eerder onderzoek volgde reeds dat TXC (taai) slecht voorspelbaar was. Daarbij bleek de taaiheid zoals voorspeld door productexperts wezenlijk te verschillen van de taaiheid voorspeld door het analytisch sensorisch panel. Ook in dit onderzoek naar statistische relaties blijkt dat TXC op basis van de eerste vijf verwerkingsronden slecht te voorspellen is. Bij de voorspelling van de vijf genoemde textuuraspecten blijken de leeftijd, het oogstjaar en het ras als extra verklarende variabele in sommige gevallen (vooral TXA en smeuïgheid) een duidelijk toegevoegde waarde te hebben. Gebruik hiervan doet echter afbreuk aan de algemene bruikbaarheid van het model en daarom zullen dergelijke extra verklarende variabelen niet zonder meer in het model worden opgenomen. Daarnaast moet het aantal parameters in een praktisch toepasbaar model minimaal zijn alsmede eenvoudig meetbaar.

De textuur van frites lijkt op basis van dit statistisch onderzoek matig te voorspellen te zijn aan de hand van grondstofeigenschappen. Mogelijke oorzaken hiervoor zijn:

- de beperkte nauwkeurigheid waarmee de textuureigenschappen kunnen worden vastgesteld;
- andere - niet gemeten of niet in de statistische verwerking meegenomen - grondstofeigenschappen zijn van belang, bijvoorbeeld bepaalde celwandeigenschappen;
- belangrijke interacties tussen proces en product;
- heterogeniteit van aardappelen en/of spreiding tussen aardappelen vertroebelt de relaties. Dit speelt tevens op produktniveau.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

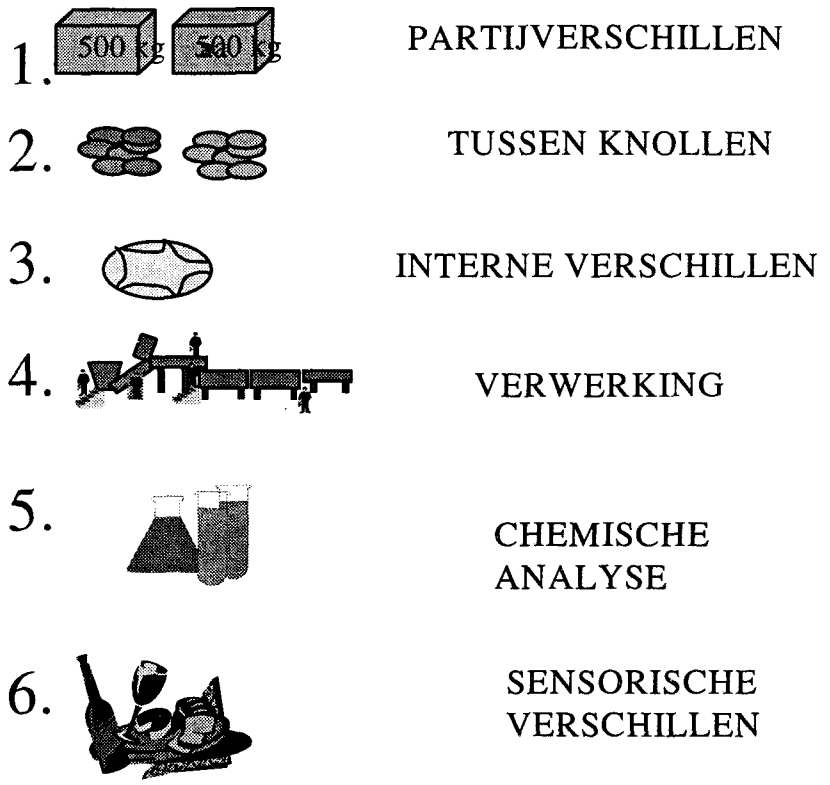
4. Discussie

Er zijn zeer veel meetmethoden ontwikkeld voor zowel het verse materiaal als aan het eindproduct. Uit de analyses is met name de spreiding in de inhoudstoffen en de variatie van de kwaliteit van het eindproduct duidelijk naar voren gekomen als problemen voor kwaliteitsbeheersing. De spreiding in kwaliteit op zich is in feite al een kwaliteitskenmerk. Als resultaat van het onderzoek is het nu mogelijk om voor elke grondstofeigenschap en voor elke productkwaliteitsanalyse aan te geven wat de spreiding is en hoe daar mee om gegaan moet worden. D.w.z. dat het bijv. duidelijk is hoe en met welke meetmethode een bepaalde textuureigenschap het best kan worden beoordeeld. En of dat dan beter door één expert drie maal gedaan moet worden of door 3 experts slechts één maal.
Het is gebleken dat sommige produkteigenschappen niet goed te meten waren, danwel omdat de meting op zich een probleem kan zijn of omdat door zeer grote spreiding in het product een bepaalde waarde weinig zegt over een partij; het is hierdoor niet te gebruiken om partijen van elkaar te onderscheiden. Door inzicht in de meetfout, spreiding in het produkt en het discriminerend vermogen van de meetmethode is het nu toch mogelijk om produktkwaliteit zo optimaal mogelijk te kwantificeren.

De strategie om aan de hand van statistische relaties nauwkeurige modellen op te stellen om daarmee producteigenschappen te voorspellen werkt niet voor alle 3 de produkteigenschappen (kleur, grauw, textuur) even goed. De vraag rijst waarom de relaties niet beter tot hun recht komen. Daar zijn verschillende redenen voor aan te geven.

Heterogeniteit
Een van de grootste bottlenecks is de enorme spreiding in grondstof tussen partijen, tussen aardappelen, en binnen één en dezelfde aardappel. Dit kan de oorzaak zijn van het uitmiddelen van effekten. Figuur 15 geeft een overzicht van de variatiebronnen.

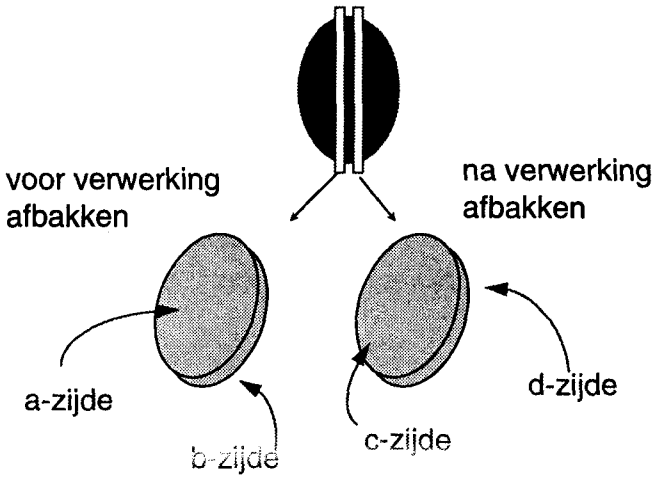
Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



Figuur 15: Bronnen van variatie

Monster-variatie
De proefopzet in het VAVI onderzoek was zodanig dat het monster dat gebruikt is voor de grondstof-samenstelling niet dezelfde is als het monster waaraan de eindkwaliteit is bepaald, alhoewel de twee monsters natuurlijk wel uit dezelfde partij komen. Het is mogelijk dat dit een fout introduceert waardoor statistische relaties niet meer goed zichtbaar zijn.
Dit probleem is onderkend en er is een opzet gekozen (bij het bakkleuronderzoek) waarbij met aardappelschijven gewerkt wordt. Het onderstaande schema (figuur 16) geeft de opzet weer. De essentie is dat twee naastliggende plakken uit één knol worden gebruikt, waarbij aan de ene de grondstof wordt geanalyseerd en aan het andere de productkwaliteit. Voor de analyse van de relatie reducerende suikers en kleur is dit grotendeels uitgewerkt. Hiermee is de variatie tussen twee monsters (of tussen knollen) omzeild en is er alleen nog de variatie binnen één knol overgebleven. Resultaten worden vermeld in het verslag Achtergrondsonderzoek.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.



b- en c-zijde zijn elkaars spiegelbeeld

Figuur 16: Door naast elkaar liggende plakken te vergelijken op grondstofsamenstelling en produkteigenschappen is het mogelijk rechtstreekse relatie te leggen. Door gebruik van deze methode worden veel bronnen van variatie geminimaliseerd.

Meetmethoden
Er zijn weliswaar veel grondstofeigenschappen bepaald maar het kan zijn dat er eigenschappen zijn die niet gemeten zijn maar die wel degelijk textuur of grauwverkleuring beïnvloeden. Een voorbeeld zou kunnen zijn specifieke celwandonderdelen waarvan het denkbaar is dat ze een invloed hebben op de textuur.

Procesinvloeden
Het proces kan een invloed hebben op de relatie grondstof-productkwaliteit. Voor de bakkleur is door een gecombineerde meting van glucose en de bakkleurindex een goede voorspelling te geven over de te verwachten eindkleur (KLA). Hierbij moet wel aangetekend worden dat deze relatie geldt voor de gekozen receptuur.
Daarnaast kan het zo zijn dat er een interactie is van de grondstof met het verwerkingsproces die er voor zorgt dat er (nog) slechts matig voorspelbare relaties worden waargenomen.
Voor andere recepten kunnen andere factoren een bijdrage leveren aan de voorspelling. Een voorbeeld hiervan is bijvoorbeeld de selectie van een blancheerproces waarbij enzymatische versteviging benut wordt.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Bijlage 1 Toelichting statistische analyse

Voor een overzicht van de correlaties werd veelal gebruik gemaakt van PCA-plaatjes (PCA=principal components analyse) in plaats van de correlatie-matrix. Een correlatie-matrix van 60 variabelen levert ongeveer 3500 aparte correlaties, terwijl een PCA een twee-dimensioneel plaatje levert waar toch de belangrijkste correlaties naar voren komen.
Voor de voorspellingsmodellen werden regressie-modellen ontwikkeld. Voor de eerste rondes zijn enkelvoudige regressies uitgevoerd met ieder van de rond 60 grondstofeigenschappen. Ook zijn regressies met twee verklarende variabelen gefit waarbij elke mogelijke combinatie van twee grondstofeigenschappen wordt genomen. Regressie op deze manier met meer dan twee variabelen was niet haalbaar vanwege het groot aantal eigenschappen - drie variabelen zou al 200.000 mogelijke combinaties leveren. Hoewel de regressies met een en twee verklarende variabelen waardevolle informaties leverden over de belangrijkste grondstofeigenschappen, werd toch verwacht dat het meenemen van enkele extra eigenschappen de verklarende waarde van het model aanzienlijk zou verbeteren. Vandaar dat na de 4e ronde gekozen is voor een andere aanpak. Een beperkte lijst van grondstofeigenschappen werd opgezet voor elk van de producteigenschap-categorieën, gebaseerd op de resultaten tot en met de derde ronde, maar ook met inbreng van expert kennis. Doordat het aantal grondstofeigenschappen nu veel kleiner was, was het ontwikkelen van modellen met 3 à 5 eigenschappen nu haalbaar. Deze aanpak wordt ook in dit eindverslag gehanteerd.

Modellen
Voor het ontwikkelen van regressie-modellen voor de voorspelling van de producteigenschappen kleur, grauw en textuur, wordt een standaard procedure gehanteerd. Deze wordt hieronder beschreven.

Procedure

Stap 1. Verklarende variabelen kiezen
Stap 1.1 Subset van 30 variabelen bepalen
Eerst wordt een lijst van 30 grondstofvariabelen gekozen voor elk van de friteseigenschappen kleur, grauw en textuur. Men beperkt zich tot die eigenschappen die in alle 5 rondes zijn gemeten. Dan nog zijn er 57 grondstofeigenschappen. Een keuze uit deze grondstofeigenschappen wordt gemaakt op basis van
i). Statistische resultaten. Concreet wordt gebruik gemaakt van de scores die voor elk grondstofeigenschap zijn berekend na de 3e verwerking. Deze scores zijn op hun beurt gebaseerd op de nauwkeurigheid van de voorspelling bij regressies van een of twee verklarende variabelen inclusief de desbetreffende eigenschap. Een eigenschap met een hoge score gaf dus een goede voorspelling.
ii). Expert kennis. Als eigenschappen volgens de experts belangrijk zouden moeten zijn, worden ze aan de lijst toegevoegd. Als voor eigenschappen op de statistische lijst volgens de experts geen zinnige verklaring van een effect kan zijn, dan worden die eigenschappen geschrapt. Als meer dan 30 eigenschappen op de lijst komen te staan, dan kiest de expert welke eigenschappen moeten vervallen.
iii). Praktijkmetingen. Een model dat (mede) gebruik maakt van metingen die sowieso in de praktijk worden uitgevoerd verdient de voorkeur. Daarom wordt onderwatergewicht

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

steeds meegenomen in plaats van %drogestof, waarmee het sterk gecorreleerd is (0,953). Ook wordt, indien relevant, de frieskleurindex altijd meegenomen.
iv). Onderlinge correlaties. Als twee variabelen sterk gecorreleerd zijn (met een correlatie van boven 0,95) dan wordt maar één van de twee meegenomen. Zo wordt voor de kleur alleen glucose en niet fructose meegenomen - de correlatie tussen de twee is 0,983.

Stap 1.2 Subset van 4 of minder variabelen kiezen
Vervolgens worden met de 30 variabelen all-possible-subsets regressies uitgevoerd. Dit houdt in dat alle mogelijke combinaties van 1,2...30 variabelen worden gefit. Als uitvoer krijgt men een lijst van de 10 beste modellen (dwz de modellen met het hoogste percentage verklaarde variantie). Ook kan men een lijst opvragen van de beste modellen met niet meer dan bijvoorbeeld 5 verklarende variabelen.
Op basis van deze resultaten wordt een subset uitgekozen van niet meer dan 4 verklarende variabelen. De resultaten laten tevens zijn hoe groot het verschil is tussen het beste model ongeacht het aantal eigenschappen en het gekozen model met 4 variabelen.

Stap 1.3 Ontwikkelen en fitten van het model
In stap 1.2 zijn alleen lineaire effecten meegenomen. Er wordt nu getoetst of kwadratische en interactie-effecten een significante verbetering geven. Zo ja, dan worden ze eveneens meegenomen in het uiteindelijk model.
Het gekozen model wordt nu gefit aan de hele dataset. De resultaten worden elders in dit verslag weergegeven. Ook wordt een uitleg gegeven van de relevante statistische maten (zie volgende pagina).

Stap 2 Bepalen voorspellingswaarde model

Het regressie-model wordt uiteindelijk gebruikt om de friteseigenschappen van onbekende partijen te voorspellen. De verklarende waarde van een model voor de bestaande dataset zegt lang niet alles over de kwaliteit van de voorspelling. Zo is het bekend dat modellen met veel verklarende variabelen weliswaar een verbetering van het percentage verklaarde variantie kunnen geven, maar doorgaans tot een slechtere voorspelling leiden.
In deze stap wordt de voorspellingswaarde beoordeeld. Het liefst zou men de voorspellingswaarde willen toetsen door het model te valideren met een nieuwe test-dataset. Die is er niet, dus wordt leave-1-out kruisvalidatie gebruikt. Een uitleg van deze techniek en de statistische maten wordt elders in volgende paragrafen gegeven.

Stap 3 Bepalen robuustheid model

Een model waarvoor de regressie-coëfficienten elk seizoen anders zijn of waar de invloed van een bepaalde eigenschap ras-afhankelijk is, heeft voor de praktijk weinig waarde. Men heeft alleen baat bij een model dat men te allen tijde kan gebruiken. In deze stap wordt dus nagegaan wat de invloed is van een drietal factoren op het model, namelijk knol-leeftijd (vers t.o.v. bewaard), oogstjaar ('93, '94, '95) en ras (Agria en Bintje). Er wordt gekeken of het model robuust is t.o.v. die factoren: blijven de geschatte coëfficienten hetzelfde, is de voorspelling net zo nauwkeurig voor verse als voor bewaarde aardappels?

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Toelichting statistische methodes

Regressiemodellen

Maten voor verklarende waarde
Er worden drie statistische maten vermeld die relevant zijn voor de beoordeling van de kwaliteit van het model.
R^2_{adj} oftewel het percentage verklaarde variantie. Hoe hoger het percentage hoe beter het regressiemodel, maar deze maat is wel dataset-afhankelijk daar het gerelateerd is aan de initiële variatie in de responsvariabel. Een bepaald regressie-model kan dus een even goede fit geven voor twee datasets, maar als de eerste dataset veel extreme objecten bevat (bijvoorbeeld uitsluitend partijen Nicola en Saturna), dan zal de R2adj voor deze dataset altijd hoger zijn. Een vergelijking van de R2adj heeft dus alleen zin als het gaat om verschillende modellen voor dezelfde dataset.
S.e.obs oftewel de standard error of observations. Deze maat zegt iets over de spreiding rond de door het regressiemodel voorspelde waarde. Hoe kleiner de s.e.obs, hoe beter het model. Een vuistregel luidt dat 95% van alle objekten (partijen dus) liggen binnen de range *voorspelde waarde +/- 2*s.e.obs.* De s.e.obs is dataset-onafhankelijk, maar moet wel gerelateerd worden aan de mogelijke range van de respons-variabelen. Zo zou een s.e.obs van 0,5 voor TXC, met een range van 0 t/m 2, veel slechter zijn dan voor GRW, met een range van 0 t/m 6.
S.e. oftewel standard error van de regressie-coëfficienten. Deze maat geeft de nauwkeurigheid aan van de geschatte regressiecoëfficient en moet klein zijn t.o.v. de coëfficient. Als de geschatte coëfficient meer dan twee keer zo groot is als de s.e. dan is er sprake van een significant effect van die variabel (de t-waarde is gelijk aan coëfficient/s.e.). Een waarschuwing: als twee verklarende variabelen met elkaar gecorreleerd zijn, dan kunnen ze allebei van invloed zijn maar toch geen significante t-waarde hebben.

Maten voor voorspellende waarde
Het doel van de statistische modellen is om de friteskwaliteit van toekomstige partijen te kunnen voorspellen. Het is dan belangrijk om niet alleen een indruk te krijgen van de verklarende waarde van een model voor de huidige dataset, maar ook van hoe goed het model een onbekende partij zou voorspellen. Dit wordt gedaan door middel van *kruisvalidatie.* Bij kruisvalidatie wordt steeds één partij weggelaten en het model gefit aan de resterende 78 partijen. Vervolgens worden de geschatte regressie-coëfficienten van het gefitte model gebruikt om de respons-waarde voor het ene weggelaten partij te voorspellen. Deze procedure wordt 79 keer herhaald waardoor elk van de 79 partijen één keer wordt weggelaten.
Kruisvalidatie levert twee maten op:
bias. Dit is het gemiddeld verschil tussen de werkelijke en voorspelde waarde per partij. De bias moet zo klein mogelijk zijn, het liefst 0.
RMSEP oftewel root mean square error of prediction. Dit is de wortel van het gemiddeld kwadraatverschil tussen werkelijke en voorspelde waarde. Deze maat moet ook zo klein mogelijk zijn. Een vuistregel luidt dat een 95% betrouwbaarheidsinterval voor de werkelijke partij-waarde wordt gegeven door voorspelde waarde +/- 2*RMSEP.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Invloed factoren (oogstjaar, ras, knolleeftijd) op regressiemodel
Men wil het liefst een model dat altijd geldig is, ongeacht het oogstjaar, het ras en periode in het bewaarseizoen. De statistische modellen konden hierop getoetst worden omdat over meerdere jaren aan meerdere rassen werd gemeten en aan zowel bewaarde als verse knollen. Tabel 13 geeft een overzicht van het aantal partijen per factor-waarde. De factoren oogstjaar en knolleeftijd zijn enigszins verstrengeld omdat voor oogstjaar 1993 alleen aan bewaarde aardappels werd gemeten. Dit betekent dat een sterk bewaareffect zich ook kan uiten als een jaareffect (en andersom). Verder zijn er slechts twee rassen (Agria en Bintje) waar voldoende vaak aan werd gemeten voor statistische toetsen.

Tabel 13: Aantal partijen per combinatie

Oogstjaar	1993	20
	1994	30
	1995	29
knol-leeftijd	vers	40
	bewaard	39
ras	Agria	15
	Bintje	16

Voor alle drie factoren is de statistische aanpak grotendeels hetzelfde.
Eerst wordt per factor-waarde een regressie uitgevoerd - dus één regressie voor verse aardappels en één regressie voor bewaarde aardappels. De geschatte regressie-coëfficienten voor de verschillende regressies worden vergeleken. Een verschil tussen de coëfficienten voor een bepaalde grondstofvariabele kan duiden op een verschil in de invloed van die variabele (bijvoorbeeld bij verse en bij bewaarde aardappels) - maar het hoeft niet. Een verschil kan ook optreden als bijvoorbeeld bij verse aardappels weinig tot geen variatie is in die grondstofvariabele. Ook kan het verschil een gevolg van toeval zijn, wat vooral het geval is als er maar weinig objecten zijn (zoals bij Agria en Bintje).
Of er sprake is van een daadwerkelijk verschillend model wordt nagegaan in de tweede stap. Hier wordt één regressie toegepast aan de gegevens van alle partijen maar met het factor als mede-verklarende variabel. De regressie-coëfficienten voor de andere verklarende variabelen mogen dan per factor-waarde varieëren. Door dit model te vergelijken met het algemeen model zonder factor kan men beoordelen of een uitsplitsing per factor-waarde een wezenlijke verbetering geeft. Met de F-toets kan men zien of er een significante verbetering is, terwijl een vergelijking van de R2adj en s.e.obs geeft aan of er sprake is van een grote verbetering.
Opmerking: voor het toetsen van het raseffect wordt steeds gebruik gemaakt van een subset bestaande uit alleen de rassen Agria en Bintje.

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

Bijlage 2 Variantie-analyse van de gemeten grondstof- en producteigen-schappen

De tabel geeft een overzicht van alle gemeten eigenschappen en de spreiding in meting en in waarden tussen de partijen.
variantie: een maat voor de spreiding
se/gemiddelde: percentage fout (maat voor de nauwkeurigheid van de bepaling)
se/range: discriminerend vermogen (maat voor verschillen tussen de partijen)
range: hoogste min de laagste waarde (van de betreffende eigenschap)

		variantie schatting partij-gemiddelde			
code	beschrijving	variantie	se/gemiddelde	se/range	range
			%	%	
grondstof					
i101[1]	OWG ongeschild	12.45	0.9	1.6	214
I102[1]	ds vers materiaal	0.049265	1.0	1.9	11.78
i103	soortelijk gewicht tussen aardappels				
I103[1]	gemiddelde	0.000005312	0.2	4.5	0.051
i103[2]	variantie	0.000000008	10.8	8.2	0.001116
i103[3]	%knollen < 1.07	0.948	21.0	1.1	90.8
i103[4]	%knollen >1.11	27.94	69.3	5.7	92.5
i104	soortelijk gewicht binnen aardappels				
i104[1]	gemiddelde	0.00004968	0.7	13.8	0.051
i104[2]	variantie	0.000000018	7.7	9.4	0.001422
i104[3]	%knol <1.045	1.981	37.6	4.9	28.8
i104[4]	%knol >1.115	2.578	33.3	3.1	52.37
i105[11]	NMR ta	5.53	1.7	1.6	143.02
i105[12]	NMR tb	83.575	1.9	1.4	653.4
i105[13]	NMR Sa	0.006355	4.4	3.2	2.529
i105[14]	NMR Sb	0.004997	3.0	1.1	6.246
i107	zetmeel	60.54	5.5	6.3	124.1
i109[1]	pH titratie	0.000000105	2.4	2.5	0.013
i109[2]	pH buffer	0.3426	2.5	2.5	23.63
i109[3]	pH midpoint pot	0.001252	0.5	3.1	1.13
i109[4]	pH zuurgraad	0.0009048	0.5	3.2	0.95
i110[1]	Na+ cation	0.001417	11.3	8.1	0.464
i110[2]	K+ cation	6.683	39.7	30.7	8.428
i110[3]	Ca++ cation	0.002209	13.4	9.6	0.491
i110[5]	Fe {new (oven)}	0.000001097	22.7	20.9	0.005001
i110[6]	Cl- anion	0.001007	7.9	2.9	1.111
i110[8]	PO4--- anion	0.008574	11.2	8.2	1.132
i110[9]	SO4-- anion	0.01289	11.4	14.0	0.81
i112[1]	oxaalzuur	0.0001583	5.5	3.8	0.328

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

code	beschrijving	variantie	se/gemiddelde	se/range	range
i112[2]	citroenzuur	0.05031	4.3	5.5	4.068
i112[3]	appelzuur	0.004531	6.7	5.0	1.3415
i112[4]	fumaarzuur	0.0001939	14.7	8.7	0.1606
i112[5]	pyroglutaminezuur	0.0001942	15.1	3.4	0.4112
i113	kleurindex CBA	0.02345	4.1	2.6	6
i113	kleurindex experts	0.05395	6.4	3.9	6
i114	glucose	0.0183	6.7	1.3	10.805
i115	fructose	0.01639	7.2	1.5	8.827
i116	sucrose	0.02067	7.2	5.0	2.865
i117[3]	Tot phenolen	0.0003733	4.4	4.0	0.4806
i118[1]	chlorogeenzuur	0.00001054	11.1	3.0	0.108
i118[2]	Vit C	0.000000091	1.0	0.4	0.0774
i118[3]	org zuur	0.06137	3.7	6.1	4.092
i119[1]	vrije aminozuren	7.094	36.6	19.5	13.67
i119[2]	eiwit (oplosbaar)	0.85	13.4	5.3	17.31
i119[3]	totaal N	0.02162	4.4	3.4	4.3
i120	dip pyro conc=0	0.4464	68.1	11.1	6
i120[2]	helling	0.00009626	48.9	8.4	0.11667
i120[3]	helling/beginwaarde	0.0001829	54.2	27.0	0.05
i121	AIS%	0.4631	0.8	5.8	11.8
i201[1]	celgrootte	10.11	1.6	4.1	78
i208[1]	maatsortering gemid (<70)	0.7214	1.5	4.6	18.56
i208[2]	maatsortering % >70	1.583	46.7	2.2	58.1
i210[1]	afkookcijfer	0.05635	11.5	5.9	4
i210[2]	grauw gekookte	0.142425	45.7	7.5	5
i212[1]	T100	0.19282	6.7	2.6	16.6
i213[1]	DSCdH	0.00708	2.2	1.2	7.18
i213[2]	DSC onset	0.040355	0.3	0.3	70.8
i215[1]	kleur chips	0.1265	9.9	5.1	7
i215[2]	textuur chips	0.0339625	6.2	8.7	2.124
i300[1]	POD	0.000000104	7.8	3.1	0.01027
i301[2]	PPO	0.00000143	1.0	0.1	0.9572
i302	PE	0.000000576	4.9	2.1	0.0355
i303	protease	0.5496	3.9	3.0	25.1
i400	%schilverlies	0.2351	3.7	6.0	8.08
i401	%sorteerverlies	0.8239	15.9	21.6	4.2
i403	%gew.verl droger	2.166666667	19.2	23.7	6.2
i404	%gew.verl koeler	0.02724	7.9	5.3	3.11
i405	%gew.verl vriezer	0.00644	n.v.t.	4.2	1.9
voorgebakken frites					
i500	%ds	0.18605	1.3	3.3	13.21
i502	%porositeit	1.28	10.3	8.0	14.12

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.

code	beschrijving	variantie	se/gemiddelde	se/range	range
i505	%vet	0.035535	4.1	9.7	1.95
nagebakken frites					
i600[1]	%ds	1.0363	2.0	6.4	15.84
i603[1]	BL blaasjes	0.01674	n.v.t.	8.3	1.5625
i603[2]	TXA nat/droog	0.125025	n.v.t.	8.8	4
i603[3]	TXB slap/hard	0.151775	n.v.t.	9.7	4
i603[4]	TXB2 hol	0.0669625	n.v.t.	34.5	0.75
i603[5]	TXC taai	0.097575	n.v.t.	20.8	1.5
i603[6]	TXD gaar	0.0089275	n.v.t.	12.6	0.75
i603[7]	SMK smaak	0.5659	n.v.t.	15.0	5
i603[8]	TXE heterogeniteit	0.138375	n.v.t.	18.6	2
i603[9]	FE feathering - allen nullen	n.v.t.	n.v.t.	n.v.t.	n.v.t.
i604[11]	krokante korst	0.030846667	n.v.t.	7.6	2.3098
i604[12]	taaie korst	0.01643	n.v.t.	8.5	1.5078
i604[13]	stevige korst	0.023555	n.v.t.	6.9	2.2383
i604[31]	dikke korst	0.017498333	n.v.t.	6.5	2.0441
i604[32]	melige totale friet	0.02052	n.v.t.	10.7	1.3407
i604[33]	droge totale friet	0.02263	n.v.t.	8.0	1.8814
i604[34]	vettig totale friet	0.023736667	n.v.t.	9.3	1.6523
i604[35]	smeuig totale friet	0.021155	n.v.t.	6.8	2.1503
i604[36]	heterogeen	0.023573333	n.v.t.	20.4	0.7508
i604[37]	stevig totale friet	0.016386667	n.v.t.	5.6	2.2723
i607[1]	GRW grauw	0.17935	n.v.t.	7.1	6
i607[2]	KLA kleur	0.11809925	n.v.t.	8.6	4
i607[4]	KLC grauw nagebakken	0.0401775	n.v.t.	4.5	4.5
i607[6]	KLB1 heter. tussen friets	0.111625	n.v.t.	16.7	2
i607[7]	KLB2 heter. binnen friets	0.1519425	n.v.t.	19.5	2
i609	%vet	0.28695	5.38	11.8	4.55

Eigendom van VAVI en ATO-DLO. Niets uit dit document mag worden gebruikt, vermeerderd of gedistribueerd zonder schriftelijke toestemming van de VAVI.